(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 September 2001 (27.09.2001)

PCT

(10) International Publication Number
WO 01/70980 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/435, 14/705, 16/18, C12N 5/10, C12Q 1/68, G01N 33/50

(21) International Application Number: PCT/US01/09341

(22) International Filing Date: 23 March 2001 (23.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/191,638	23 March 2000 (23.03.2000)	US
09/618,893	18 July 2000 (18.07.2000)	US

(71) Applicant: PE CORPORATION (NY) [US/US]; Robert A. Millman, Assistant Secretary, 761 Main Avenue, Norwalk, CT 06859 (US).

(72) Inventor: CRAVCHIK, Anibal; Celera, 45 West Gude Drive, Rockville, MD 20850 (US).

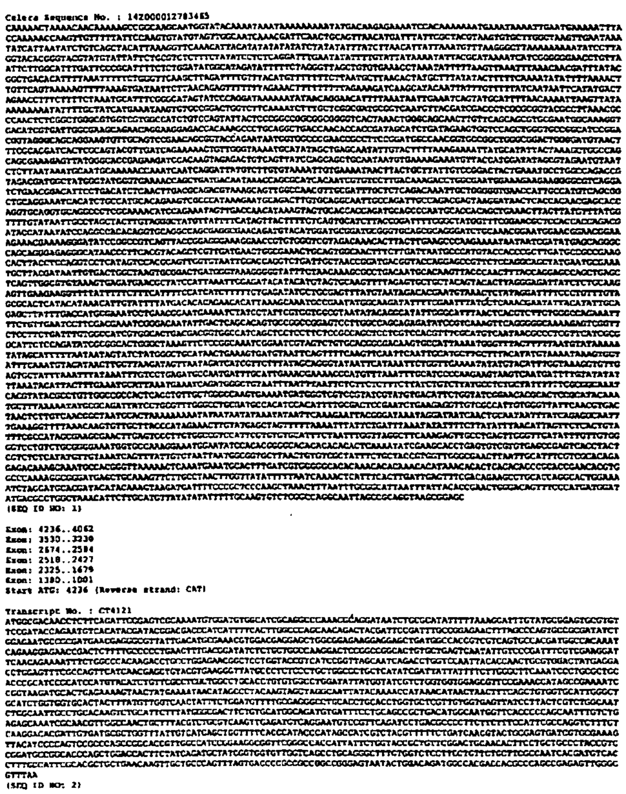
(74) Agent: MILLMAN, Robert, A.; Celera Genomics Corp., 45 West Gude Drive C2-4, Rockville, MD 20850 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: ISOLATED G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES THEREOF AS INSECTICIDAL TARGETS



(57) Abstract: The present invention provides amino acid sequences of proteins that are encoded by genes within the *Drosophila melanogaster* genome, the GPCR proteins of the present invention. The present invention specifically provides isolated protein and nucleic acid molecules, methods of identifying orthologs and paralogs of the GPCR proteins and methods of identifying modulators of the GPCR proteins for use as insecticides.

WO 01/70980 A3

Published:
- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report: 18 April 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/09341

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/435 C07K14/705 C07K16/18 C12N5/10 C12Q1/68 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K G01N C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

SEQUENCE SEARCH, EPO-Internal, BIOSIS, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL SEQUENCE LIBRARY [Online] 6 January 2000 (2000-01-06) ADAMS, M. AND VENTER, J.C.: "Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces" XP002182628 accession no AC020076 ---	1,2,4,5
A	WO 94 08006 A (ZYMOGENETICS INC) 14 April 1994 (1994-04-14) the whole document ---	

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 November 2001	15.02.2002
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Holtorf, S

2

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LI X-J ET AL: "CLONING, FUNCTIONAL EXPRESSION, AND DEVELOPMENTAL REGULATION OF A NEUROPEPTIDE Y RECEPTOR FROM DROSOPHILA MELANOGASTER" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 267, no. 1, 5 January 1992 (1992-01-05), pages 9-12, XP000877443 ISSN: 0021-9258 the whole document ---	
A	WO 99 01468 A (DEN HEUVEL MARCEL VAN ;INGHAM PHILIP W (GB); ONTOGENY INC (US)) 14 January 1999 (1999-01-14) the whole document ---	
A	HAUSER FRANK ET AL: "Molecular cloning, genomic organization and developmental regulation of a novel receptor from Drosophila melanogaster structurally related to gonadotropin-releasing hormone receptors from vertebrates." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 249, no. 3, 28 August 1998 (1998-08-28), pages 822-828, XP002182627 ISSN: 0006-291X the whole document ---	
A	FENG G ET AL: "CLONING AND FUNCTIONAL CHARACTERIZATION OF A NOVEL DOPAMINE RECEPTOR FROM DROSOPHILA MELANOGASTER" JOURNAL OF NEUROSCIENCE, NEW YORK, NY, US, vol. 15, no. 12, 15 June 1995 (1995-06-15), pages 3925-3933, XP002919142 ISSN: 0270-6474 the whole document ---	
P,X	ADAMS M D ET AL: "THE GENOME SEQUENCE OF DROSOPHILA MELANOGASTER" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 287, no. 5461, 24 March 2000 (2000-03-24), pages 2185-2195, XP000961051 ISSN: 0036-8075 the whole document ---	1,2,4,5

-/--

2

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE EMBL SEQUENCE LIBRARY [Online] 24 March 2000 (2000-03-24) ADAMS, M.D:, ETAL.: "Drosophila melanogaster genomic scaffold 142000013386053 section 8 of 30 complete sequence" XP002182629 accession no. AE003491 and AE002593 ---	1,2,4,5
P,X	DATABASE TREMBLREL. DATABASE [Online] 1 May 2000 (2000-05-01) ADAMS, M.D., ET AL.: "The genome sequence of Drosophila melanogaster - cross-reference to EMBL accession no. AE003491, AAF48216.1" XP002182630 accession no. Q9VYH9 ---	1,2
E	EP 1 136 501 A (BAYER AG) 26 September 2001 (2001-09-26) the whole document -----	1-20

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/09341

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-20 partially

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: claims 1-20 partially

Isolated protein consisting or comprising an amino acid sequence as characterized by SEQID3, or an allelic variant or an ortholog of said amino acid sequence wherein said variant or ortholog is encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule as characterized by SEQID1 or 2; an antibody that binds to said protein; furthermore a nucleic acid molecule consisting or comprising of a nucleotide sequence that
1) encodes the amino acid sequence of SEQId3
2) encodes and allelic variant or an ortholog of an amino acid sequence of SEQId3 wherein said nucleotide sequence hybridizes to SEQID 1 or 2
3) encodes a fragment of said SEQID3
4) is the complement of the nucleotides of 1) to 3)

The recombinant expression of the same in host cells and methods for the detection of said proteins or said nucleic acids in a sample with the help of an agent that binds to said protein or an oligonucleotide and kits that contain such agent or oligonucleotide.
Furthermore, a method to identify an agent that binds to said protein by detecting a complex formed by an agent and the said protein.

Invention 2-66: claims 1-20 partially

as invention one but referring to the protein and nucleic acid sequences as characterized by SEQIDs 6,9,...,192,195,198; SEQIDs 4,7,....,190,193,196 and SEQIDs 5,8,....,191,194,197, respectively.

''TERNATIONAL SEARCH REPORT

Information on patent family members

I .national Application No
PCT/US 01/09341

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 9408006	A	14-04-1994	CA	2145866	A1	14-04-1994
			EP	0663006	A1	19-07-1995
			JP	8501942	T	05-03-1996
			WO	9408006	A1	14-04-1994
			US	5683884	A	04-11-1997
			US	5674981	A	07-10-1997
			US	5622839	A	22-04-1997
			US	5674689	A	07-10-1997
WO 9901468	A	14-01-1999	AU	8380898	A	25-01-1999
			WO	9901468	A2	14-01-1999
EP 1136501	A	26-09-2001	DE	10013618	A1	20-09-2001
			EP	1136501	A2	26-09-2001
			JP	2001299369	A	30-10-2001

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 September 2001 (27.09.2001)

PCT

(10) International Publication Number
WO 01/070980 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/435, 14/705, 16/18, C12N 5/10, C12Q 1/68, G01N 33/50

(21) International Application Number: PCT/US01/09341

(22) International Filing Date: 23 March 2001 (23.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/191,638 23 March 2000 (23.03.2000) US
09/618,893 18 July 2000 (18.07.2000) US

(71) Applicant: PE CORPORATION (NY) [US/US]; Robert A. Millman, Assistant Secretary, 761 Main Avenue, Norwalk, CT 06859 (US).

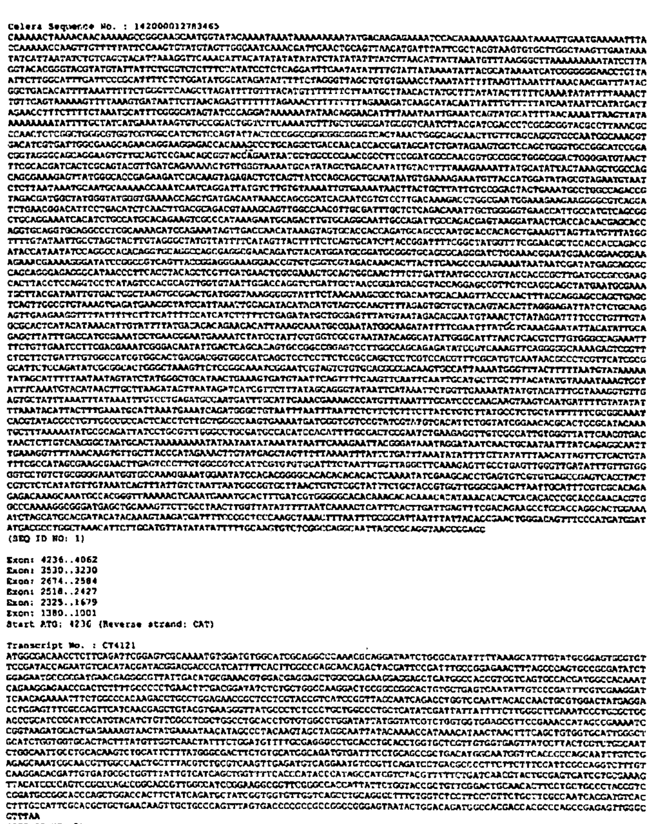
(72) Inventor: CRAVCHIK, Anibal; Celera, 45 West Gude Drive, Rockville, MD 20850 (US).

(74) Agent: MILLMAN, Robert, A.; Celera Genomics Corp., 45 West Gude Drive C2-4, Rockville, MD 20850 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: ISOLATED G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES THEREOF AS INSECTICIDAL TARGETS



(57) Abstract: The present invention provides amino acid sequences of proteins that are encoded by genes within the *Drosophila melanogaster* genome, the GPCR proteins of the present invention. The present invention specifically provides isolated protein and nucleic acid molecules, methods of identifying orthologs and paralogs of the GPCR proteins and methods of identifying modulators of the GPCR proteins for use as insecticides.

WO 01/070980 A3

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) **Date of publication of the international search report:** 18 April 2002

(48) **Date of publication of this corrected version:** 16 January 2003

(15) **Information about Correction:**
see PCT Gazette No. 03/2003 of 16 January 2003, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

ISOLATED G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES THEREOF AS INSECTICIDAL TARGETS

RELATED APPLICATIONS

The present application claims priority to U.S. Serial No. 60/191,638, filed March 23, 2000 (Atty. Docket CL000388) and U.S. Serial No. 09/618,893, filed July 18, 2000 (Atty. Docket CL000733).

FIELD OF THE INVENTION

The present invention is in the field of G-Protein coupled receptors (GPCRs), recombinant DNA molecules and protein production. The present invention specifically provides novel GPCR proteins and nucleic acid molecules encoding such protein molecules, for use in the development of insecticide and insecticide targets and as a source for identifying human therapeutics and human therapeutic development.

BACKGROUND OF THE INVENTION

The *Drosophila melanogaster* genome is 165 Mb, with about 120 Mb of this being euchromatic. The genome is organized in 4 chromosome pairs and is estimated to contain 10 - 12,000 genes. Model organisms, such as *Drosophila melanogaster,* share many genes with humans whose sequences and functions have been conserved. In addition to myriad similarities in cellular structure and function, humans and Drosophila share pathways for intercellular signaling, developmental patterning, learning and behavior, as well as tumor formation and metastasis.

The genes involved in the development of Drosophila, with few exceptions, are the same as those involved in the development of higher organisms. Developmental biology studies the sequential activation and interaction of genes, in relation to developing morphology. Right now, Drosophila is the only organism for which one can begin with a list of genes active in the egg and follow the morphological changes and gene activations through to adulthood.

Drosophila studies have provided the widest knowledge base available for any single organism; accordingly, developmental biologists use the fly to ferret out the activity of genes with similar functions in higher organisms. Despite its small size, the fly is by no means a small developmental problem. If you know the genes involved in the development of the fly, you also

know, to a reasonable approximation, the genes involved in the development of the worm, the fish, the mouse, and humans.

A major goal in drug/insecticide development is to understand and elucidate the molecular mechanisms that govern cell signaling and cell-cell interactions in higher eukaryotes. GPCRs form major links in cellular communication/response systems. A complete list of GPCR proteins from Drosophila would therefore be invaluable in developing human therapeutic targets. Not only will the proteins serve as models for human cellular signaling and response, such molecules will also serve as molecular keys in identifying therapeutically important human orthologs.

G-protein coupled receptors

G-protein coupled receptors (GPCRs) constitute a major class of proteins responsible for transducing a signal within a cell. GPCRs have three structural domains: an amino terminal extracellular domain, a transmembrane domain containing seven transmembrane segments, three extracellular loops, and three intracellular loops, and a carboxy terminal intracellular domain. Upon binding of a ligand to an extracellular portion of a GPCR, a signal is transduced within the cell that results in a change in a biological or physiological property of the cell. GPCRs, along with G-proteins and effectors (intracellular enzymes and channels modulated by G-proteins), are the components of a modular signaling system that connects the state of intracellular second messengers to extracellular inputs.

GPCR genes and gene-products are potential causative agents of disease (Spiegel *et al.*, *J. Clin. Invest.* *92*:1119-1125 (1993); McKusick *et al.*, *J. Med. Genet.* *30*:1-26 (1993)). Specific defects in the rhodopsin gene and the V2 vasopressin receptor gene have been shown to cause various forms of retinitis pigmentosum (Nathans *et al.*, *Annu. Rev. Genet.* *26*:403-424(1992)), and nephrogenic diabetes insipidus (Holtzman *et al.*, *Hum. Mol. Genet.* *2*:1201-1204 (1993)). These receptors are of critical importance to both the central nervous system and peripheral physiological processes. Evolutionary analyses suggest that the ancestor of these proteins originally developed in concert with complex body plans and nervous systems.

The GPCR protein superfamily can be divided into five families: Family I, receptors typified by rhodopsin and the β2-adrenergic receptor and currently represented by over 200 unique members (Dohlman *et al.*, *Annu. Rev. Biochem.* *60*:653-688 (1991)); Family II, the parathyroid hormone/calcitonin/secretin receptor family (Juppner *et al.*, *Science* *254*:1024-1026 (1991); Lin *et al.*, *Science* *254*:1022-1024 (1991)); Family III, the metabotropic glutamate receptor family

(Nakanishi, *Science 258* 597:603 (1992)); Family IV, the cAMP receptor family, important in the chemotaxis and development of *D. discoideum* (Klein *et al.*, *Science 241*:1467-1472 (1988)); and Family V, the fungal mating pheromone receptors such as STE2 (Kurjan, *Annu. Rev. Biochem. 61*:1097-1129 (1992)).

There are also a small number of other proteins, which present seven putative hydrophobic segments and appear to be unrelated to GPCRs; they have not been shown to couple to G-proteins. *Drosophila* expresses a photoreceptor-specific protein, bride of sevenless (boss), a seven-transmembrane-segment protein, which has been extensively studied and does not show evidence of being a GPCR (Hart *et al.*, *Proc. Natl. Acad. Sci. USA 90*:5047-5051 (1993)). The gene *frizzled* (*fz*) in *Drosophila* is also thought to be a protein with seven transmembrane segments. Like boss, fz has not been shown to couple to G-proteins (Vinson *et al.*, *Nature 338*:263-264 (1989)).

G proteins represent a family of heterotrimeric proteins composed of α, β and γ subunits, that bind guanine nucleotides. These proteins are usually linked to cell surface receptors, e.g., receptors containing seven transmembrane segments. Following ligand binding to the GPCR, a conformational change is transmitted to the G protein, which causes the α-subunit to exchange a bound GDP molecule for a GTP molecule and to dissociate from the βγ-subunits. The GTP-bound form of the α-subunit typically functions as an effector-modulating moiety, leading to the production of second messengers, such as cAMP (e.g., by activation of adenyl cyclase), diacylglycerol or inositol phosphates. Greater than 20 different types of α-subunits are known in humans. These subunits associate with a smaller pool of β and γ subunits. Examples of mammalian G proteins include Gi, Go, Gq, Gs and Gt. G proteins are described extensively in Lodish *et al.*, *Molecular Cell Biology*, (Scientific American Books Inc., New York, N.Y., 1995), the contents of which are incorporated herein by reference. GPCRs, G proteins and G protein-linked effector and second messenger systems have been reviewed in *The G-Protein Linked Receptor Fact Book*, Watson *et al.*, eds., Academic Press (1994).

<u>Dopamine receptors</u>

The understanding of the dopaminergic system relevance in brain function and disease developed several decades ago from three diverse observations following drug treatments. These were the observations that dopamine replacement therapy improved Parkinson's disease symptoms, depletion of dopamine and other catecholamines by reserpine caused depression and antipsychotic drugs blocked dopamine receptors. The finding that the dopamine receptor binding affinities of typical antipsychotic drugs correlate with their clinical potency led to the dopamine overactivity

hypothesis of schizophrenia (Snyder, S.H., *Am J Psychiatry 133*, 197-202 (1976); Seeman, P. and Lee, T., *Science 188*, 1217-9 (1975)). Today, dopamine receptors are crucial targets in the pharmacological therapy of schizophrenia, Parkinson's disease, Tourette's syndrome, tardive dyskinesia and Huntington's disease. The dopaminergic system includes the nigrostriatal, mesocorticolimbic and tuberoinfundibular pathways. The nigrostriatal pathway is part of the striatal motor system and its degeneration leads to Parkinson's disease; the mesocorticolimbic pathway plays a key role in reinforcement and in emotional expression and is the desired site of action of antipsychotic drugs; the tuberoinfundibular pathways regulates prolactin secretion from the pituitary.

Dopamine receptors are members of the G protein coupled receptor superfamily, a large group proteins that share a seven helical membrane-spanning structure and transduce signals through coupling to heterotrimeric guanine nucleotide-binding regulatory proteins (G proteins). Dopamine receptors are classified into subfamilies: D1-like (D1 and D5) and D2-like (D2, D3 and D4) based on their different ligand binding profiles, signal transduction properties, sequence homologies and genomic organizations (Civelli, O., Bunzow, J.R. and Grandy, D.K., *Annu Rev Pharmacol Toxicol 33*, 281-307 (1993)). The D1-like receptors, D1 and D5, stimulate cAMP synthesis through coupling with Gs-like proteins and their genes do not contain introns within their protein coding regions. On the other hand, the D2-like receptors, D2, D3 and D4, inhibit cAMP synthesis through their interaction with Gi-like proteins and share a similar genomic organization which includes introns within their protein coding regions.

Serotonin receptors

Serotonin (5-Hydroxytryptamine; 5-HT) was first isolated from blood serum, where it was shown to promote vasoconstriction (Rapport, M.M., Green, A.A. and Page, I.H., *J Biol Chem 176*, 1243-1251 (1948). Interest on a possible relationship between 5-HT and psychiatric disease was spurred by the observations that hallucinogens such as LSD and psilocybin inhibit the actions of 5-HT on smooth muscle preparations (Gaddum, J.H. and Hameed, K.A., *Br J Pharmacol 9*, 240-248 (1954)). This observation lead to the hypothesis that brain 5-HT activity might be altered in psychiatric disorders (Wooley, D.W. and Shaw, E., *Proc Natl Acad Sci U S A 40*, 228-231 (1954); Gaddum, J.H. and Picarelli, Z.P., *Br J Pharmacol 12*, 323-328 (1957)). This hypothesis was strengthened by the introduction of tricyclic antidepressants and monoamine oxidase inhibitors for the treatment of major depression and the observation that those drugs affected noradrenaline and 5-HT metabolism. Today, drugs acting on the serotoninergic system have been proved to be effective

in the pharmacotherapy of psychiatric diseases such as depression, schizophrenia, obsessive-compulsive disorder, panic disorder, generalized anxiety disorder and social phobia as well as migraine, vomiting induced by cancer chemotherapy and gastric motility disorders.

Serotonin receptors represent a very large and diverse family of neurotransmitter receptors. To date thirteen 5-HT receptor proteins coupled to G proteins plus one ligand–gated ion channel receptor (5-HT3) have been described in mammals. This receptor diversity is thought to reflect serotonin's ancient origin as a neurotransmitter and a hormone as well as the many different roles of 5-HT in mammals. The 5-HT receptors have been classified into seven subfamilies or groups according to their different ligand-binding affinity profiles, molecular structure and intracellular transduction mechanisms (Hoyer, D. et al., *Pharmacol. Rev. 46*, 157-203 (1994)).

GPCRs are a major target for drug/insecticide action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize previously unknown GPCRs. The present invention advances the state of the art by providing a previously unidentified drosophila GPCRs.

Insecticides

About 10,000 species of the more than 1 million species of insects are crop-eating, and of these, approximately 700 species worldwide cause most of the insect damage to man's crops, in the field and in storage.

A detailed study of novel proteins from Drosophila and invertebrate orthologs thereof, will serve as targets for identifying new members of the known classes of insecticides as well as aiding in the identification of new classes of compounds.

SUMMARY OF THE INVENTION

The present invention is based in part on the identification of nucleic acid sequences that encode amino acid sequences of 66 GPCR proteins that are produced by *Drosophila melanogaster* and invertebrate, human and other mammalian orthologs thereof. These unique protein sequences, and nucleic acid sequences that encode these proteins, can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic and insecticidal agents.

The proteins of the present inventions are GPCRs that participate in signaling pathways. As used herein, a "signaling pathway" refers to the modulation (e.g., stimulation or inhibition) of a cellular function/activity upon the binding of a ligand to the GPCR protein. Examples of such

functions include mobilization of intracellular molecules that participate in a signal transduction pathway, e.g., phosphatidylinositol 4,5-bisphosphate (PIP_2), inositol 1,4,5-triphosphate (IP_3) and adenylate cyclase; polarization of the plasma membrane; production or secretion of molecules; alteration in the structure of a cellular component; cell proliferation, e.g., synthesis of DNA; cell migration; cell differentiation; and cell survival

The response mediated by the receptor protein depends on the type of cell. For example, in some cells, binding of a ligand to the receptor protein may stimulate an activity such as release of compounds, gating of a channel, cellular adhesion, migration, differentiation, etc., through phosphatidylinositol or cyclic AMP metabolism and turnover while in other cells, the binding of the ligand will produce a different result. Regardless of the cellular activity/response modulated by the receptor protein, it is universal that the protein is a GPCR and interacts with G proteins to produce one or more secondary signals, in a variety of intracellular signal transduction pathways, e.g., through phosphatidylinositol or cyclic AMP metabolism and turnover, in a cell.

As used herein, "phosphatidylinositol turnover and metabolism" refers to the molecules involved in the turnover and metabolism of phosphatidylinositol 4,5-bisphosphate (PIP_2) as well as to the activities of these molecules. PIP_2 is a phospholipid found in the cytosolic leaflet of the plasma membrane. Binding of ligand to the receptor activates, in some cells, the plasma-membrane enzyme phospholipase C that in turn can hydrolyze PIP_2 to produce 1,2-diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP_3). Once formed IP_3 can diffuse to the endoplasmic reticulum surface where it can bind an IP_3 receptor, e.g., a calcium channel protein containing an IP_3 binding site. IP_3 binding can induce opening of the channel, allowing calcium ions to be released into the cytoplasm. IP_3 can also be phosphorylated by a specific kinase to form inositol 1,3,4,5-tetraphosphate (IP_4), a molecule that can cause calcium entry into the cytoplasm from the extracellular medium. IP_3 and IP_4 can subsequently be hydrolyzed very rapidly to the inactive products inositol 1,4-biphosphate (IP_2) and inositol 1,3,4-triphosphate, respectively. These inactive products can be recycled by the cell to synthesize PIP_2. The other second messenger produced by the hydrolysis of PIP_2, namely 1,2-diacylglycerol (DAG), remains in the cell membrane where it can serve to activate the enzyme protein kinase C. Protein kinase C is usually found soluble in the cytoplasm of the cell, but upon an increase in the intracellular calcium concentration, this enzyme can move to the plasma membrane where it can be activated by DAG. The activation of protein kinase C in different cells results in various cellular responses such as the phosphorylation of glycogen synthase, or the phosphorylation of various transcription factors, e.g., NF-kB. The language "phosphatidylinositol activity", as used herein, refers to an activity of PIP_2 or one of its metabolites.

Another signaling pathway in which the receptor may participate is the cAMP turnover pathway. As used herein, "cyclic AMP turnover and metabolism" refers to the molecules involved in the turnover and metabolism of cyclic AMP (cAMP) as well as to the activities of these molecules. Cyclic AMP is a second messenger produced in response to ligand-induced stimulation of certain G protein coupled receptors. In the cAMP signaling pathway, binding of a ligand to a GPCR can lead to the activation of the enzyme adenyl cyclase, which catalyzes the synthesis of cAMP. The newly synthesized cAMP can in turn activate a cAMP-dependent protein kinase. This activated kinase can phosphorylate a voltage-gated potassium channel protein, or an associated protein, and lead to the inability of the potassium channel to open during an action potential. The inability of the potassium channel to open results in a decrease in the outward flow of potassium, which normally repolarizes the membrane of a neuron, leading to prolonged membrane depolarization.

By targeting an agent to modulate a GPCR, the signaling activity and biological process mediated by the receptor can be agonized or antagonized. Such agonism and antagonism serves as a basis for modulating a biological activity in a therapeutic context (mammalian therapy) or toxic context (anti-invertebrate/insecticidal agent).

DESCRIPTION OF THE FIGURE SHEETS

FIGURE SHEETS 1-89 provides genomic nucleic acid sequences from Drosophila melanogaster, predicted transcript, amino acid coding regions, information relating to the subfamily of GPCR proteins to which the protein sequence belongs, e.g., BLAST hit homology, presence of hidden Markov model, etc. Also provided are 5' promoter sequences, transciption initiation sites and other structural features.

DETAILED DESCRIPTION OF THE INVENTION

General Description

The present invention is based on the sequencing of the *Drosophila melanogaster* genome. During the sequencing and assembly of the *Drosophila melanogaster* genome, analysis of the sequence information revealed previously unidentified nucleic acid molecules that encode proteins that share structural and/or sequence homology to protein/peptide/domains identified and characterized within the art as being a GPCR protein or part of a GPCR protein. Based on this analysis, the present invention provides amino acid sequences of 66 GPCR proteins that are produced by Drosophila melanogaster, nucleic acid sequences that encode these GPCR proteins,

and the art known protein/peptide/domain that have structural or sequence homology to the Drosophila melanogaster protein.

In addition to being previously unknown, the proteins and nucleic acid molecules that are provided in the present invention are selected based on their ability to be used for the development of commercially important products and services. Specifically, the present proteins are selected based on homology and/or structural relatedness to known GPCR proteins of commercial importance and/or the presence of genetic signals that define the protein as a GPCR protein. Some of the more specific features of the proteins of the present invention, and the uses thereof, are described in detail below, and are known within the art for each category of GPCR.

Specific Embodiments

Peptide Molecules

In the Figure sheets, the present invention provides nucleic acid sequences that encode 66 protein molecules that have been identified as being within the Drosophila melanogaster genome (SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, encoded by nucleic acid molecules provided in SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)). These protein sequences will be referred herein as the GPCR proteins of the present invention, GPCR proteins, or the peptides or the proteins of the present invention. Table 1 provides a summary of the GPCRs of the present invention and the subfamily assignment of each one.

The present invention provides isolated peptide and protein molecules that consist of, consist essentially of or are comprised of the amino acid sequences of the GPCR proteins disclosed in the Figure sheets (SEQ ID NO: 3, 6, 9, 189, 192, 195, 198, encoded by nucleic acid molecules provided in SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)), as well as all obvious variants of these proteins that are within the art to make and use. Some of these variants are described in detail below.

As used herein, a protein is said to be "isolated" or "purified" when it is substantially free of cellular material or free of chemical precursors or other chemicals. The proteins of the present invention can be purified to homogeneity or other degrees of purity. The level of purification will be based on the intended use. The critical feature is that the preparation allows for the desired function of the protein, even if in the presence of considerable amounts of other components.

In some uses, "substantially free of cellular material" includes preparations of the protein having less than about 30% (by dry weight) other proteins (i.e., contaminating protein), less than about 20% other proteins, less than about 10% other proteins, or less than about 5% other proteins. When the protein is recombinantly produced, it can also be substantially free of culture medium, i.e., culture medium represents less than about 20% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of the protein in which it is separated from chemical precursors or other chemicals that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of the GPCR protein having less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other chemicals.

The isolated GPCR protein can be purified from cells that naturally express it, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods. For example, a nucleic acid molecule encoding the GPCR protein is cloned into an expression vector, the expression vector introduced into a host cell and the protein expressed in the host cell. The protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Many of these techniques are described in detail below.

Accordingly, the present invention provides proteins that consist of one of the amino acid sequences encoded by the nucleic acid sequences shown in the Figure sheets (SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)). The amino acid sequences of such proteins are provided in the Figure Sheets along with an explanation of the protein type/family. A protein consists of an amino acid sequence when the amino acid sequence is the final amino acid sequence of the protein.

The present invention further provides proteins that consist essentially of one of the amino acid sequences encoded by the nucleic acid sequences shown in the Figure sheets (SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)). A protein consists essentially of an amino acid sequence when such an amino acid sequence is present with only a few additional amino acid residues in the final protein.

The present invention further provides proteins that are comprised of one of the amino acid sequences encoded by the nucleic acid sequences shown in the Figure sheets (SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript

sequences)). A protein is comprised of an amino acid sequence when the amino acid sequence is at least part of the final amino acid sequence of the protein. In such a fashion, the protein can be only the protein or have additional amino acid molecules, such as amino acid residues (contiguous encoded sequence) that are naturally associated with it or heterologous amino acid residues/peptide sequences. Such a protein can have a few additional amino acid residues or can comprise several hundred or more additional amino acids. The preferred classes of proteins that are comprised of the GPCR proteins of the present invention are the naturally occurring mature proteins.

The GPCR proteins of the present invention can be attached to heterologous sequences to form chimeric or fusion proteins. Such chimeric and fusion proteins comprise a GPCR protein operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the GPCR protein. "Operatively linked" indicates that the GPCR protein and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the GPCR protein.

In some uses, the fusion protein does not affect the activity of the GPCR protein *per se*. For example, the fusion protein can include, but is not limited to, enzymatic fusion proteins, for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions, MYC-tagged, HI-tagged and Ig fusions. Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant GPCR protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of a protein can be increased by using a heterologous signal sequence.

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see Ausubel *et al.*, *Current Protocols in Molecular Biology*, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST protein). A GPCR protein-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the GPCR protein.

As mentioned above, the present invention also provides and enables obvious variants of the amino acid sequence of the proteins of the present invention, such as naturally occurring mature forms of the protein, allelic/sequence variants of the proteins, non-naturally occurring

recombinantly derived variants of the proteins, and orthologs and paralogs of the proteins. Such variants can readily be generated using art know techniques in the fields of recombinant nucleic acid technology and protein biochemistry. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

Such variants can readily be identified/made using molecular techniques and the sequence information disclosed herein. Further, such variants can readily be distinguished from other proteins based on sequence and/or structural homology to the GPCR proteins of the present invention. The degree of homology/identity present will be based primarily on whether the protein is a functional variant or non-functional variant, the amount of divergence present in the paralog family and the evolutionary distance between the orthologs.

To determine the percent identity of two amino acid sequences or two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, 40%, 50%, 60%, 70%, 80%, or 90% or more of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity and similarity between two sequences can be accomplished using a mathematical algorithm. (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part 1*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm

which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux, J., *et al., Nucleic Acids Res. 12(1)*:387 (1984)) (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against sequence databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (*J. Mol. Biol.* 215:403-10 (1990)). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to the proteins of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (*Nucleic Acids Res.* 25(17):3389-3402 (1997)). When utilizing BLAST and gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

Full-length pre-processed forms, as well as mature processed forms, of proteins that comprise one of the proteins of the present invention can readily be identified as having complete sequence identity to one of the GPCR proteins of the present invention as well as being encoded by the same genetic locus as the GPCR protein provided herein.

Allelic variants of a GPCR protein can readily be identified as having a high degree (significant) of sequence homology/identity to at least a portion of the GPCR protein as well as being encoded by the same genetic locus as the GPCR protein provided herein. As used herein, two proteins (or a region of the proteins) have significant homology when the amino acid sequences are typically at least about 70-75%, 80-85%, and more typically at least about 90-95% or more homologous. A significantly homologous amino acid sequence, according to the present

invention, will be encoded by a nucleic acid sequence that will hybridize to a GPCR protein encoding nucleic acid molecule under stringent conditions as more fully described below.

Paralogs of a GPCR protein can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the GPCR protein, as being encoded by a gene from Drosophila, and as having similar activity or function. Two proteins will typically be considered paralogs when the amino acid sequences are typically at least about 70-75%, 80-85%, and more typically at least about 90-95% or more homologous through a given region or domain. Such paralogs will be encoded by a nucleic acid sequence that will hybridize to a GPCR protein encoding nucleic acid molecule under stringent conditions as more fully described below.

Orthologs of a GPCR protein can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the GPCR protein as well as being encoded by a gene from another organism. Preferred orthologs will be isolated from mammals, preferably human, for the development of human therapeutic targets and agents, or other invertebrates, particularly insects of economical/agriculture importance, e.g. members of the Lepidopteran and Coleopteran orders, for the development of insecticides and insecticidal targets. Such orthologs will be encoded by a nucleic acid sequence that will hybridize to a GPCR protein encoding nucleic acid molecule under moderate to stringent conditions, as more fully described below, depending on the degree of relatedness of the two organisms yielding the proteins.

Non-naturally occurring variants of the GPCR proteins of the present invention can readily be generated using recombinant techniques. Such variants include, but are not limited to deletions, additions and substitutions in the amino acid sequence of the GPCR protein. For example, one class of substitutions are conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a GPCR protein by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al., Science 247*:1306-1310 (1990).

Variant GPCR proteins can be fully functional or can lack function in one or more activities, e.g. ability to bind ligand, ability to bind G-protein, ability to mediate signaling, etc. Fully functional variants typically contain only conservative variation or variation in non-critical residues or in non-critical regions. Functional variants can also contain substitution of similar amino acids

that result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree.

Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.*, *Science 244*:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or in assays such as an *in vitro* proliferative activity. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al., J. Mol. Biol. 224*:899-904 (1992); de Vos *et al. Science 255*:306-312 (1992)).

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally-occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in polypeptides are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art.

Accordingly, the polypeptides also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence for purification of the mature polypeptide or a pro-protein sequence.

The present invention further provides fragments of the GPCR proteins, in addition to proteins and peptides that comprise and consist of such fragments. The fragments to which the invention pertains, however, are not to be construed as encompassing fragments that may be disclosed publicly prior to the present invention.

As used herein, a fragment comprises at least 8 or more contiguous amino acid residues from a GPCR protein. Such fragments can be chosen based on the ability to retain one or more of the biological activities of the GPCR protein or could be chosen for the ability to perform a function, e.g. act as an immunogen. Particularly important fragments are biologically active

fragments, peptides which are, for example, about 8 or more amino acids in length. Such fragments will typically comprise a domain or motif of the GPCR protein, e.g., active site, a G-protein binding site, a transmembrane domain or a ligand binding domain. Further, possible fragments include, but are not limited to, domain or motif containing fragments, soluble peptide fragments, and fragments containing immunogenic structures. Predicted domains and functional sites are readily identifiable by computer programs well-known and readily available to those of skill in the art (e.g., PROSITE analysis).

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally-occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in GPCR proteins are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art.

Accordingly, the GPCR proteins of the present invention also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature GPCR protein is fused with another compound, such as a compound to increase the half-life of the GPCR protein (for example, polyethylene glycol), or in which the additional amino acids are fused to the mature GPCR protein, such as a leader or secretory sequence or a sequence for purification of the mature GPCR protein or a pro-protein sequence.

Known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well-known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins - Structure and*

Molecular Properties, 2nd Ed., T.E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (*Meth. Enzymol. 182*: 626-646 (1990)) and Rattan *et al.* (*Ann. N.Y. Acad. Sci. 663*:48-62 (1992)).

Protein/Peptide Uses

The proteins of the present invention can be used in assays to identify modulators as potential insecticides, to determine the biological activity of the protein, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its binding partner or receptor) in biological fluids; and as markers for tissues in which the corresponding protein is preferentially expressed. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the binding partner so as to develop a system to identify inhibitors of the binding interaction. Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987 and can readily be determined using the information provided in Table 1.

The potential uses of the proteins of the present invention are based primarily on the source of the protein as well as the class/action of the protein. For example, GPCRs isolated from Drosophila and other invertebrates serve as a target for identifying anti-invertebrate compounds, e.g. insecticides. GPCRs isolated from Drosophila and their human/mammalian orthologs serve as targets for identifying agents for use in mammalian therapeutic applications, e.g. a human drug. Approximately 70% of all pharmaceutical agents modulate the activity of a GPCR. A combination of the invertebrate and mammalian ortholog can be used in selective screening methods to find agents specific for invertebrates. Table 1 provides specific context of use for each of the presently disclosed GPCRs. Particularly useful are GPCRs involved in neurotransmission for use as insecticide targets.

The receptor polypeptides (including variants and fragments which may have been disclosed prior to the present invention) are useful for biological assays related to GPCRs. Such assays involve any of the known GPCR functions or activities or properties useful for diagnosis and treatment of GPCR-related conditions.

The receptor polypeptides are also useful in drug/insecticide screening assays, in cell-based or cell-free systems. Cell-based systems can be native, i.e., cells that normally express the receptor protein, as a biopsy or expanded in cell culture. In one embodiment, however, cell-based assays involve recombinant host cells expressing the receptor protein.

The polypeptides can be used to identify compounds that modulate receptor activity. Both the GPCRs of the present invention and appropriate variants and fragments can be used in high-throughput screens to assay candidate compounds for the ability to bind to the receptor. These compounds can be further screened against a functional receptor to determine the effect of the compound on the receptor activity. Further, these compounds can be tested in animal or invertebrate systems to determine activity/effectiveness. Compounds can be identified that activate (agonist) or inactivate (antagonist) the receptor to a desired degree.

Further, the receptor polypeptides can be used to screen a compound for the ability to stimulate or inhibit interaction between the receptor protein and a target molecule that normally interacts with the receptor protein. The target can be ligand or a component of the signal pathway that the receptor protein normally interacts (for example, a G-protein or other interactor involved in cAMP or phosphatidylinositol turnover and/or adenylate cyclase, or phospholipase C activation). Such assays typically include the steps of combining the receptor protein with a candidate compound under conditions that allow the receptor protein, or fragment, to interact with the target molecule, and to detect the formation of a complex between the protein and the target or to detect the biochemical consequence of the interaction with the receptor protein and the target, such as any of the associated effects of signal transduction such as G-protein phosphorylation, cAMP or phosphatidylinositol turnover, and adenylate cyclase or phospholipase C activation.

Candidate compounds include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam *et al.*, *Nature 354*:82-84 (1991); Houghten *et al.*, *Nature 354*:84-86 (1991)) and combinatorial chemistry-derived molecular libraries made of D- and/or L- configuration amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang *et al.*, *Cell 72*:767-778 (1993)); 3) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, $F(ab')_2$, Fab expression library

fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules (e.g., molecules obtained from combinatorial and natural product libraries).

One candidate compound is a soluble fragment of the receptor that competes for ligand binding. Other candidate compounds include mutant receptors or appropriate fragments containing mutations that affect receptor function and thus compete for ligand. Accordingly, a fragment that competes for ligand, for example with a higher affinity, or a fragment that binds ligand but does not allow release, is encompassed by the invention.

The invention further includes other end point assays to identify compounds that modulate (stimulate or inhibit) receptor activity. The assays typically involve an assay of events in the signal transduction pathway that indicate receptor activity. Thus, the expression of genes that are up- or down-regulated in response to the receptor protein dependent signal cascade can be assayed. In one embodiment, the regulatory region of such genes can be operably linked to a marker that is easily detectable, such as luciferase. Alternatively, phosphorylation of the receptor protein, or a receptor protein target, could also be measured.

Any of the biological or biochemical functions mediated by the receptor can be used as an endpoint assay. These include all of the biochemical or biochemical/biological events described herein, in the references cited herein, incorporated by reference for these endpoint assay targets, and other functions known to those of ordinary skill in the art.

Binding and/or activating compounds can also be screened by using chimeric receptor proteins in which the amino terminal extracellular domain, or parts thereof, the entire transmembrane domain or subregions, such as any of the seven transmembrane segments or any of the intracellular or extracellular loops and the carboxy terminal intracellular domain, or parts thereof, can be replaced by heterologous domains or subregions. For example, a G-protein-binding region can be used that interacts with a different G-protein then that which is recognized by the native receptor. Accordingly, a different set of signal transduction components is available as an end-point assay for activation. Alternatively, the entire transmembrane portion or subregions (such as transmembrane segments or intracellular or extracellular loops) can be replaced with the entire transmembrane portion or subregions specific to a host cell that is different from the host cell from which the amino terminal extracellular domain and/or the G-protein-binding region are derived. This allows for assays to be performed in other than the specific host cell from which the receptor is derived. Alternatively, the amino terminal extracellular domain (and/or other ligand-binding regions) could be replaced by a domain (and/or other binding region) binding a different ligand, thus, providing an assay for test compounds that interact with the heterologous amino terminal

extracellular domain (or region) but still cause signal transduction. Finally, activation can be detected by a reporter gene containing an easily detectable coding region operably linked to a transcriptional regulatory sequence that is part of the native signal transduction pathway.

The receptor polypeptides are also useful in competition binding assays in methods designed to discover compounds that interact with the receptor. Thus, a compound is exposed to a receptor polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble receptor polypeptide is also added to the mixture. If the test compound interacts with the soluble receptor polypeptide, it decreases the amount of complex formed or activity from the receptor target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the receptor. Thus, the soluble polypeptide that competes with the target receptor region is designed to contain peptide sequences corresponding to the region of interest.

To perform cell free drug/insecticide screening assays, it is sometimes desirable to immobilize either the receptor protein, or fragment, or its target molecule to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

Techniques for immobilizing proteins on matrices can be used in the drug/insecticide screening assays. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase/15625 fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g., ^{35}S-labeled) and the candidate compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of receptor-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques. For example, either the polypeptide or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin using techniques well known in the art. Alternatively, antibodies reactive with the protein but which do not interfere with binding of the protein to its target molecule can be derivatized to the wells of the plate, and the protein trapped in the wells by antibody conjugation. Preparations of a receptor-binding protein and a candidate compound are incubated in the receptor protein-presenting wells and the amount of complex trapped in the well

can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the receptor protein target molecule, or which are reactive with receptor protein and compete with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

Agents that modulate one of the GPCRs of the present invention can be identified using one or more of the above assays, alone or in combination. It is generally preferable to use a cell-based or cell free system first and then confirm activity in an animal/insect model system. Such model systems are well known in the art and can readily be employed in this context.

Modulators of receptor protein activity identified according to these drug/insecticide screening assays can be used to treat a subject with a disorder mediated by the receptor pathway, by treating cells that express the GPCR. These methods of treatment include the steps of administering the modulators of protein activity in a pharmaceutical composition as described herein, to a subject in need of such treatment.

In yet another aspect of the invention, the GPCR proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Biotechniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with the GPCR and are involved in GPCR activity. Such GPCR-binding proteins are also likely to be involved in the propagation of signals by the GPCR proteins or GPCR targets as, for example, downstream elements of a GPCR-mediated signaling pathway, e.g., a signaling pathway. Alternatively, such GPCR-binding proteins are likely to be GPCR inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a GPCR protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo,* forming a GPCR-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription

factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the GPCR protein.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a GPCR modulating agent, an antisense GPCR nucleic acid molecule, a GPCR-specific antibody, or a GPCR-binding partner) can be used in an animal or insect model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal or insect model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for insecticidal activity as described herein.

Antibodies

The invention also provides antibodies that selectively bind to one of the proteins of the present invention, a protein comprising such a peptide, as well as variants and fragments thereof. As used herein, an antibody selectively binds a target peptide when it binds the target peptide and does not significantly bind to unrelated proteins. An antibody is still considered to selectively bind a peptide even if it also binds to other proteins that are not substantially homologous with the target peptide so long as such proteins share homology with a fragment or domain of the peptide target of the antibody. In this case, it would be understood that antibody binding to the peptide is still selective despite some degree of cross-reactivity.

As used herein, an antibody is defined in terms consistent with that recognized within the art: they are multi-subunit proteins produced by a mammalian organism in response to an antigen challenge. The antibodies of the present invention include polyclonal antibodies and monoclonal antibodies, as well as fragments of such antibodies, including, but not limited to, Fab or $F(ab')_2$, and Fv fragments.

Many methods are known for generating and/or identifying antibodies to a given target peptide. Several such methods are described by Harlow, Antibodies, Cold Spring Harbor Press, (1989).

In general, to generate antibodies, an isolated peptide is used as an immunogen and is administered to a mammalian organism, such as a rat, rabbit or mouse. Either the full-length protein, an antigenic peptide fragment or a fusion protein can be used.

Antibodies are preferably prepared from regions or discrete fragments of the GPCR proteins. Antibodies can be prepared from any region of the protein as described herein. However, preferred regions will include those involved in function/activity and/or receptor/binding partner interaction.

An antigenic fragment will typically comprise at least 10 contiguous amino acid residues. The antigenic peptide can comprise, however, at least 12, 14, 20 or more amino acid residues. Such fragments can be selected on a physical property, such as fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions or can be selected based on sequence uniqueness.

Detection on an antibody of the present invention can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Antibody Uses

The antibodies can be used to isolate one of the proteins of the present invention by standard techniques, such as affinity chromatography or immunoprecipitation. The antibodies can facilitate the purification of the natural protein from cells and recombinantly produced protein expressed in host cells. In addition, such antibodies are useful to detect the presence of one of the proteins of the present invention in cells or tissues to determine the pattern of expression of the protein among various tissues in an organism and over the course of normal development. Further, such antibodies can be used to detect protein *in situ*, *in vitro*, or in a cell lysate or supernatant in order to evaluate the abundance and pattern of expression. Also, such antibodies can be used to assess abnormal tissue

distribution or abnormal expression during development. Antibody detection of circulating fragments of the full length protein can be used to identify turnover.

The antibodies can also be used to assess normal and aberrant subcellular localization of cells in the various tissues in an organism. The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a treatment modality. Accordingly, where treatment is ultimately aimed at correcting expression level or the presence of aberrant sequence and aberrant tissue distribution or developmental expression, antibodies directed against the or relevant fragments can be used to monitor therapeutic efficacy.

The antibodies are also useful for inhibiting protein function, for example, blocking the binding of the GPCR protein to a binding partner such as a ligand. An antibody can be used, for example, to block binding, thus modulating (agonizing or antagonizing) the peptides activity. Antibodies can be prepared against specific fragments containing sites required for function or against intact protein that is associated with a cell or cell membrane.

The invention also encompasses kits for using antibodies to detect the presence of a protein in a biological sample. The kit can comprise antibodies such as a labeled or labelable antibody and a compound or agent for detecting protein in a biological sample; means for determining the amount of protein in the sample; means for comparing the amount of protein in the sample with a standard; and instructions for use.

<u>Nucleic Acid Molecules</u>

The present invention further provides isolated nucleic acid molecules that encode a GPCR protein or protein of the present invention (SEQ ID NO: 3, 6, 9, . . . , 192, 195, 198, encoded by nucleic acid molecules provided in SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)). Such nucleic acid molecules will consist of, consist essentially of, or comprise a nucleotide sequence that encodes one of the GPCR proteins of the present invention, an allelic variant thereof, or an ortholog or paralog thereof.

As used herein, an "isolated" nucleic acid molecule is one that is separated from other nucleic acid present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. However, there can be some flanking nucleotide sequences, for example up to about 5KB, 4KB, 3KB, 2KB, or 1KB or less, particularly contiguous protein encoding sequences and protein

encoding sequences within the same gene but separated by introns in the genomic sequence. The important point is that the nucleic acid is isolated from remote and unimportant flanking sequences such that it can be subjected to the specific manipulations described herein such as recombinant expression, preparation of probes and primers, and other uses specific to the nucleic acid sequences.

Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. However, the nucleic acid molecule can be fused to other coding or regulatory sequences and still be considered isolated.

For example, recombinant DNA molecules contained in a vector are considered isolated. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the isolated DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Accordingly, the present invention provides nucleic acid molecules that consist of one of the nucleotide sequences shown in the Figure sheets (SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)). A nucleic acid molecule consists of a nucleotide sequence when the nucleotide sequence is the complete nucleotide sequence of the nucleic acid molecule.

The present invention further provides nucleic acid molecules that consist essentially of one of the nucleotide sequences shown in the Figure sheets (SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)). A nucleic acid molecule consists essentially of a nucleotide sequence when such a nucleotide sequence is present with only a few additional nucleic acid residues in the final nucleic acid molecule.

The present invention further provides nucleic acid molecules that are comprised of one of the nucleotide sequences shown in the Figure sheets (SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)). A nucleic acid molecule is comprised of a nucleotide sequence when the nucleotide sequence is at least part of the final nucleotide sequence of the nucleic acid molecule. In such a fashion, the nucleic acid molecule can be only the nucleotide sequence or have additional nucleic acid residues, such as nucleic acid residues that are naturally associated with it or heterologous nucleotide

sequences. Such a nucleic acid molecule can have a few additional nucleotides or can comprises several hundred or more additional nucleotides. The preferred classes of nucleic acid molecules that are comprised of the nucleotide sequences of the present are the naturally occurring full-length cDNA molecules and genes and genomic clones since some of the nucleic acid molecules provided in the Figure sheets are fragments of the complete gene that exists in nature. A brief description of how various types of these nucleic acid molecules can be readily made/isolated is provided below.

In the Figures, both coding and non-coding sequences are provided for each protein encoding nucleic acid sequence. Because of the source of the present invention, Drosophila genomic sequences, the nucleic acid molecules in the figures will contain genomic intronic sequences, 5' and 3' non-coding sequences, gene regulatory regions and non-coding intergenic sequences. In general such sequence features are either noted or can readily be identified using computational tools known in the art. As discussed below, some of the non-coding regions, particularly gene regulatory elements such as promoters, are useful for a variety of purposes, e.g. control of heterologous gene expression, target for identifying gene activity modulating compounds.

Full-length genes may be cloned from known sequence using any one of a number of methods known in the art. For example, a method which employs XL-PCR (Perkin-Elmer, Foster City, Calif.) to amplify long pieces of DNA may be used. Other methods for obtaining full-length sequences are well known in the art.

The isolated nucleic acid molecules can encode the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature protein (when the mature form has more than one protein chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, facilitate protein trafficking, prolong or shorten protein half-life or facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in situ*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

As mentioned above, the isolated nucleic acid molecules include, but are not limited to, the sequence encoding the GPCR protein alone, the sequence encoding the mature protein and additional coding sequences, such as a leader or secretory sequence (e.g., a pre-pro or pro-protein sequence), the sequence encoding the mature protein, with or without the additional coding sequences, plus additional non-coding sequences, for example introns and non-coding 5' and 3' sequences such as transcribed but non-translated sequences that play a role in transcription, mRNA processing (including splicing and polyadenylation signals), ribosome binding and stability of

mRNA. In addition, the nucleic acid molecule may be fused to a marker sequence encoding, for example, a protein that facilitates purification.

Isolated nucleic acid molecules can be in the form of RNA, such as mRNA, or in the form DNA, including cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The nucleic acid, especially DNA, can be double-stranded or single-stranded. Single-stranded nucleic acid can be the coding strand (sense strand) or the non-coding strand (anti-sense strand).

The invention further provides nucleic acid molecules that encode fragments of the proteins of the present invention as well as nucleic acid molecules that encode obvious variants of the GPCR proteins of the present invention that are described above. Such nucleic acid molecules may be naturally occurring, such as allelic variants (same locus), paralogs (different locus), and orthologs (different organism), or may be constructed by recombinant DNA methods or by chemical synthesis. Such non-naturally occurring variants may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells, or organisms. Accordingly, as discussed above, the variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions.

The present invention further provides non-coding fragments of the nucleic acid molecules provided in the Figures. Preferred non-coding fragments include, but are not limited to, promoter sequences, enhancer sequences, gene modulating sequences and gene termination sequences. Such fragments are useful in controlling heterologous gene expression and in developing screens to identify gene modulating agents. Particularly useful are fragments from about 100 to about 1,000 nucleotides taken 5' from the start ATG in the genomic sequences provided in the Figures (SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196,).

A fragment comprises a contiguous nucleotide sequence greater than 12 or more nucleotides. Further, a fragment could at least 30, 40, 50, 100, 250 or 500 nucleotides in length. The length of the fragment will be based on its intended use. For example, the fragment can encode epitope bearing regions of the protein, or can be useful as DNA probes and primers. Such fragments can be isolated using the known nucleotide sequence to synthesize an oligonucleotide probe. A labeled probe can then be used to screen a cDNA library, genomic DNA library, or mRNA to isolate nucleic acid corresponding to the coding region. Further, primers can be used in PCR reactions to clone specific regions of gene.

A probe/primer typically comprises substantially a purified oligonucleotide or oligonucleotide pair. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 20, 25, 40, 50 or more consecutive nucleotides.

Orthologs, homologs, and allelic variants can be identified using methods well known in the art. As described in the protein Section, these variants comprise a nucleotide sequence encoding a protein that is typically 60-65%, 70-75%, 80-85%, and more typically at least about 90-95% or more homologous to the nucleotide sequence shown in the Figure sheets or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under moderate to stringent conditions, to the nucleotide sequence shown in the Figure sheets or a fragment of the sequence.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences encoding a protein at least 50-55% homologous to each other typically remain hybridized to each other. The conditions can be such that sequences at least about 65%, at least about 70%, or at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. One example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65C.

Nucleic Acid Molecule Uses

The nucleic acid molecules of the present invention are useful for probes, primers, chemical intermediates, and in biological assays. The nucleic acid molecules are useful as a hybridization probe for cDNA and genomic DNA to isolate full-length cDNA and genomic clones encoding the protein described in the Figures and to isolate cDNA and genomic clones that correspond to variants (alleles, orthologs, etc.) producing the same or related proteins shown in the Figures.

The probe can correspond to any sequence along the entire length of the nucleic acid molecules provided in the Figures. Accordingly, it could be derived from 5' noncoding regions, the coding region, and 3' noncoding regions. However, as discussed, fragments are not to be construed as encompassing fragments disclosed prior to the present invention.

The nucleic acid molecules are also useful as primers for PCR to amplify any given region of a nucleic acid molecule and are useful to synthesize antisense molecules of desired length and sequence.

The nucleic acid molecules are also useful for constructing recombinant vectors. Such vectors include expression vectors that express a portion of, or all of, the protein sequences. Vectors also include insertion vectors, used to integrate into another nucleic acid molecule sequence, such as into the cellular genome, to alter *in situ* expression of a gene and/or gene product. For example, an endogenous coding sequence can be replaced via homologous recombination with all or part of the coding region containing one or more specifically introduced mutations.

The nucleic acid molecules are also useful for expressing antigenic portions of the proteins.

The nucleic acid molecules are also useful in making vectors containing the gene regulatory regions of the nucleic acid molecules of the present invention.

The nucleic acid molecules are also useful for designing ribozymes corresponding to all, or a part, of the mRNA produced from the nucleic acid molecules described herein.

The nucleic acid molecules are also useful for constructing host cells expressing a part, or all, of the nucleic acid molecules and proteins.

The nucleic acid molecules are also useful for constructing transgenic animals expressing all, or a part, of the nucleic acid molecules and proteins.

The nucleic acid molecules are also useful for making vectors that express part, or all, of the proteins.

The nucleic acid molecules are also useful as hybridization probes for determining the presence, level, form and distribution of nucleic acid expression. Accordingly, the probes can be used to detect the presence of, or to determine levels of, a specific nucleic acid molecule in cells, tissues, and in organisms. The nucleic acid whose level is determined can be DNA or RNA. Accordingly, probes corresponding to the proteins described herein can be used to assess expression and/or gene copy number in a given cell, tissue, or organism. These uses are relevant for efficacy of insecticides involving an increase or decrease in GPCR protein expression relative to normal results.

In vitro techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detecting DNA includes Southern hybridizations and *in situ* hybridization.

Probes can be used as a part of a diagnostic test kit for identifying cells or tissues that express a GPCR protein, such as by measuring a level of a receptor-encoding nucleic acid in a

sample of cells from a subject e.g., mRNA or genomic DNA, or determining if a receptor gene has been mutated.

Nucleic acid expression assays are useful for drug/insecticide screening to identify compounds that modulate GPCR nucleic acid expression.

The invention thus provides a method for identifying a compound that can be used to treat a block the activity of the GPCR by modulating the expression of the GPCR gene. The method typically includes assaying the ability of the compound to modulate the expression of the GPCR nucleic acid and thus identifying a compound that can be used to kill insects by altering GPCR nucleic acid expression. The assays can be performed in cell-based and cell-free systems. Cell-based assays include cells naturally expressing the GPCR nucleic acid or recombinant cells genetically engineered to express specific nucleic acid sequences.

The assay for GPCR nucleic acid expression can involve direct assay of nucleic acid levels, such as mRNA levels, or on collateral compounds involved in the signal pathway. Further, the expression of genes that are up- or down-regulated in response to the GPCR protein signal pathway can also be assayed. In this embodiment the regulatory regions of these genes can be operably linked to a reporter gene such as luciferase.

Thus, modulators of GPCR gene expression can be identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of GPCR mRNA in the presence of the candidate compound is compared to the level of expression of GPCR mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder characterized by aberrant nucleic acid expression. When expression of mRNA is statistically significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of nucleic acid expression. When nucleic acid expression is statistically significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of nucleic acid expression.

The nucleic acid molecules are also useful in diagnostic assays for qualitative changes in GPCR nucleic acid, and particularly in qualitative changes that lead to insecticidal activity/tolerance. The nucleic acid molecules can be used to detect mutations in GPCR genes and gene expression products such as mRNA.

The nucleic acid molecules are thus useful as antisense constructs to control GPCR gene expression in cells, tissues, and organisms. A DNA antisense nucleic acid molecule is designed to

be complementary to a region of the gene involved in transcription, preventing transcription and hence production of GPCR protein. An antisense RNA or DNA nucleic acid molecule would hybridize to the mRNA and thus block translation of mRNA into GPCR protein.

The invention also encompasses kits for detecting the presence of a GPCR nucleic acid in a biological sample. For example, the kit can comprise reagents such as a labeled or labelable nucleic acid or agent capable of detecting GPCR nucleic acid in a biological sample; means for determining the amount of GPCR nucleic acid in the sample; and means for comparing the amount of GPCR nucleic acid in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect GPCR protein mRNA or DNA.

Vectors/host cells

The invention also provides vectors containing the nucleic acid molecules described herein. The term "vector" refers to a vehicle, preferably a nucleic acid molecule, that can transport the nucleic acid molecules. When the vector is a nucleic acid molecule, the nucleic acid molecules are covalently linked to the vector nucleic acid. With this aspect of the invention, the vector includes a plasmid, single or double stranded phage, a single or double stranded RNA or DNA viral vector, or artificial chromosome, such as a BAC, PAC, YAC, OR MAC.

A vector can be maintained in the host cell as an extrachromosomal element where it replicates and produces additional copies of the nucleic acid molecules. Alternatively, the vector may integrate into the host cell genome and produce additional copies of the nucleic acid molecules when the host cell replicates.

The invention provides vectors for the maintenance (cloning vectors) or vectors for expression (expression vectors) of the nucleic acid molecules. The vectors can function in procaryotic or eukaryotic cells or in both (shuttle vectors).

Expression vectors contain cis-acting regulatory regions that are operably linked in the vector to the nucleic acid molecules such that transcription of the nucleic acid molecules is allowed in a host cell. The nucleic acid molecules can be introduced into the host cell with a separate nucleic acid molecule capable of affecting transcription. Thus, the second nucleic acid molecule may provide a trans-acting factor interacting with the cis-regulatory control region to allow transcription of the nucleic acid molecules from the vector. Alternatively, a trans-acting factor may be supplied by the host cell. Finally, a trans-acting factor can be produced from the vector itself. It

is understood, however, that in some embodiments, transcription and/or translation of the nucleic acid molecules can occur in a cell-free system.

The regulatory sequence to which the nucleic acid molecules described herein can be operably linked include promoters for directing mRNA transcription. These include, but are not limited to, the left promoter from bacteriophage λ, the lac, TRP, and TAC promoters from *E. coli*, the early and late promoters from SV40, the CMV immediate early promoter, the adenovirus early and late promoters, and retrovirus long-terminal repeats.

In addition to control regions that promote transcription, expression vectors may also include regions that modulate transcription, such as repressor binding sites and enhancers. Examples include the SV40 enhancer, the cytomegalovirus immediate early enhancer, polyoma enhancer, adenovirus enhancers, and retrovirus LTR enhancers.

In addition to containing sites for transcription initiation and control, expression vectors can also contain sequences necessary for transcription termination and, in the transcribed region a ribosome binding site for translation. Other regulatory control elements for expression include initiation and termination codons as well as polyadenylation signals. The person of ordinary skill in the art would be aware of the numerous regulatory sequences that are useful in expression vectors. Such regulatory sequences are described, for example, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

A variety of expression vectors can be used to express a nucleic acid molecule. Such vectors include chromosomal, episomal, and virus-derived vectors, for example vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, including yeast artificial chromosomes, from viruses such as baculoviruses, papovaviruses such as SV40, Vaccinia viruses, adenoviruses, poxviruses, pseudorabies viruses, and retroviruses. Vectors may also be derived from combinations of these sources such as those derived from plasmid and bacteriophage genetic elements, e.g. cosmids and phagemids. Appropriate cloning and expression vectors for prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual. 2nd. ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

The regulatory sequence may provide constitutive expression in one or more host cells (i.e. tissue specific) or may provide for inducible expression in one or more cell types such as by temperature, nutrient additive, or exogenous factor such as a hormone or other ligand. A variety of

vectors providing for constitutive and inducible expression in prokaryotic and eukaryotic hosts are well known to those of ordinary skill in the art.

The nucleic acid molecules can be inserted into the vector nucleic acid by well-known methodology. Generally, the DNA sequence that will ultimately be expressed is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction enzymes and then ligating the fragments together. Procedures for restriction enzyme digestion and ligation are well known to those of ordinary skill in the art.

The vector containing the appropriate nucleic acid molecule can be introduced into an appropriate host cell for propagation or expression using well-known techniques. Bacterial cells include, but are not limited to, *E. coli*, *Streptomyces*, and *Salmonella typhimurium*. Eukaryotic cells include, but are not limited to, yeast, insect cells such as *Drosophila*, animal cells such as COS and CHO cells, and plant cells.

As described herein, it may be desirable to express the protein as a fusion protein. Accordingly, the invention provides fusion vectors that allow for the production of the proteins. Fusion vectors can increase the expression of a recombinant protein, increase the solubility of the recombinant protein, and aid in the purification of the protein by acting for example as a ligand for affinity purification. A proteolytic cleavage site may be introduced at the junction of the fusion moiety so that the desired peptide can ultimately be separated from the fusion moiety. Proteolytic enzymes include, but are not limited to, factor Xa, thrombin, and enterokinase. Typical fusion expression vectors include pGEX (Smith *et al.*, *Gene 67*:31-40 (1988)), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, *Gene 69*:301-315 (1988)) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology 185*:60-89 (1990)).

Recombinant protein expression can be maximized in a host bacteria by providing a genetic background wherein the host cell has an impaired capacity to proteolytically cleave the recombinant protein. (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Alternatively, the sequence of the nucleic acid molecule of interest can be altered to provide preferential codon usage for a specific host cell, for example *E. coli*. (Wada *et al.*, *Nucleic Acids Res. 20*:2111-2118 (1992)).

The nucleic acid molecules can also be expressed by expression vectors that are operative in yeast. Examples of vectors for expression in yeast e.g., *S. cerevisiae* include pYepSec1 (Baldari, *et*

al., EMBO J. 6:229-234 (1987)), pMFa (Kurjan *et al., Cell 30*:933-943(1982)), pJRY88 (Schultz *et al., Gene 54*:113-123 (1987)), and pYES2 (Invitrogen Corporation, San Diego, CA).

The nucleic acid molecules can also be expressed in insect cells using, for example, baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith *et al., Mol. Cell Biol. 3*:2156-2165 (1983)) and the pVL series (Lucklow *et al., Virology 170*:31-39 (1989)).

In certain embodiments of the invention, the nucleic acid molecules described herein are expressed in mammalian cells using mammalian expression vectors. Examples of mammalian expression vectors include pCDM8 (Seed, B. *Nature 329*:840(1987)) and pMT2PC (Kaufman *et al., EMBO J. 6*:187-195 (1987)).

The expression vectors listed herein are provided by way of example only of the well-known vectors available to those of ordinary skill in the art that would be useful to express the nucleic acid molecules. The person of ordinary skill in the art would be aware of other vectors suitable for maintenance propagation or expression of the nucleic acid molecules described herein. These are found for example in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

The invention also encompasses vectors in which the nucleic acid sequences described herein are cloned into the vector in reverse orientation, but operably linked to a regulatory sequence that permits transcription of antisense RNA. Thus, an antisense transcript can be produced to all, or to a portion, of the nucleic acid molecule sequences described herein, including both coding and non-coding regions. Expression of this antisense RNA is subject to each of the parameters described above in relation to expression of the sense RNA (regulatory sequences, constitutive or inducible expression, tissue-specific expression).

The invention also relates to recombinant host cells containing the vectors described herein. Host cells therefore include prokaryotic cells, lower eukaryotic cells such as yeast, other eukaryotic cells such as insect cells, and higher eukaryotic cells such as mammalian cells.

The recombinant host cells are prepared by introducing the vector constructs described herein into the cells by techniques readily available to the person of ordinary skill in the art. These include, but are not limited to, calcium phosphate transfection, DEAE-dextran-mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, lipofection, and other techniques such as those found in Sambrook, *et al.* (*Molecular Cloning: A*

Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Host cells can contain more than one vector. Thus, different nucleotide sequences can be introduced on different vectors of the same cell. Similarly, the nucleic acid molecules can be introduced either alone or with other nucleic acid molecules that are not related to the nucleic acid molecules such as those providing trans-acting factors for expression vectors. When more than one vector is introduced into a cell, the vectors can be introduced independently, co-introduced or joined to the nucleic acid molecule vector.

In the case of bacteriophage and viral vectors, these can be introduced into cells as packaged or encapsulated virus by standard procedures for infection and transduction. Viral vectors can be replication-competent or replication-defective. In the case in which viral replication is defective, replication will occur in host cells providing functions that complement the defects.

Vectors generally include selectable markers that enable the selection of the subpopulation of cells that contain the recombinant vector constructs. The marker can be contained in the same vector that contains the nucleic acid molecules described herein or may be on a separate vector. Markers include tetracycline or ampicillin-resistance genes for prokaryotic host cells and dihydrofolate reductase or neomycin resistance for eukaryotic host cells. However, any marker that provides selection for a phenotypic trait will be effective.

While the mature proteins can be produced in bacteria, yeast, mammalian cells, and other cells under the control of the appropriate regulatory sequences, cell- free transcription and translation systems can also be used to produce these proteins using RNA derived from the DNA constructs described herein.

Where secretion of the protein is desired, which is difficult to achieve with multi-transmembrane domain containing proteins such as GPCRs, appropriate secretion signals are incorporated into the vector. The signal sequence can be endogenous to the proteins or heterologous to these proteins.

Where the protein is not secreted into the medium, which is typically the case with GPCRs, the protein can be isolated from the host cell by standard disruption procedures, including freeze thaw, sonication, mechanical disruption, use of lysing agents and the like. The protein can then be recovered and purified by well-known purification methods including ammonium sulfate precipitation, acid extraction, anion or cationic exchange chromatography, phosphocellulose chromatography, hydrophobic-interaction chromatography, affinity chromatography,

hydroxylapatite chromatography, lectin chromatography, or high performance liquid chromatography.

It is also understood that depending upon the host cell in recombinant production of the proteins described herein, the proteins can have various glycosylation patterns, depending upon the cell, or maybe non-glycosylated as when produced in bacteria. In addition, the proteins may include an initial modified methionine in some cases as a result of a host-mediated process.

Nucleic Acid Arrays

The present invention further provides arrays or microarrays of nucleic acid molecules that are based on the sequence information provided in the Figure Sheets SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)).

As used herein "Arrays" or "Microarrays" refers to an array of distinct polynucleotides or oligonucleotides synthesized on a substrate, such as paper, nylon or other type of membrane, filter, chip, glass slide, or any other suitable solid support. In one embodiment, the microarray is prepared and used according to the methods described in US Patent 5,837,832, Chee et al., PCT application W095/11995 (Chee et al.), Lockhart, D. J. et al. (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. et al. (1996; Proc. Natl. Acad. Sci. 93: 10614-10619), all of which are incorporated herein in their entirety by reference. In other embodiments, such arrays are produced by the methods described by Brown et. al., US Patent No. 5,807,522.

The microarray is preferably composed of a large number of unique, single-stranded nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6-60 nucleotides in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25 nucleotides in length. For a certain type of microarray, it may be preferable to use oligonucleotides that are only 7-20 nucleotides in length. The microarray may contain oligonucleotides that cover the known 5', or 3', sequence, sequential oligonucleotides which cover the full length sequence; or unique oligonucleotides selected from particular areas along the length of the sequence. Polynucleotides used in the microarray may be oligonucleotides that are specific to a gene or genes of interest.

In order to produce oligonucleotides to a known sequence for a microarray, the gene(s) of interest (or an ORF identified from the contigs of the present invention) is typically examined using a computer algorithm which starts at the 5' or at the 3' end of the nucleotide sequence.

Typical algorithms will then identify oligomers of defined length that are unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that may interfere with hybridization. In certain situations it may be appropriate to use pairs of oligonucleotides on a microarray. The "pairs" will be identical, except for one nucleotide that preferably is located in the center of the sequence. The second oligonucleotide in the pair (mismatched by one) serves as a control. The number of oligonucleotide pairs may range from two to one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support.

In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application W095/251116 (Baldeschweiler et al.) which is incorporated herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536, 6144 or more oligonucleotides, or any other number between two and one million which lends itself to the efficient use of commercially available instrumentation.

In order to conduct sample analysis using a microarray, the RNA or DNA from a biological sample is made into hybridization probes. The mRNA is isolated, and cDNA is produced and used as a template to make antisense RNA (aRNA). The aRNA is amplified in the presence of fluorescent nucleotides, and labeled probes are incubated with the microarray so that the probe sequences hybridize to complementary oligonucleotides of the microarray. Incubation conditions are adjusted so that hybridization occurs with precise complementary matches or with various degrees of less complementarity. After removal of nonhybridized probes, a scanner is used to determine the levels and patterns of fluorescence. The scanned images are examined to determine degree of complementarity and the relative abundance of each oligonucleotide sequence on the microarray. The biological samples may be obtained from any bodily fluids (such as blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. A detection system may be used to measure the absence, presence, and amount of hybridization for all of the distinct sequences simultaneously. This data may be used for large

scale correlation studies on the sequences, expression patterns, mutations, variants, or polymorphisms among samples.

Using such arrays, the present invention provides methods to identify the expression of one or more of the proteins/peptides of the present invention. In detail, such methods comprise incubating a test sample with one or more nucleic acid molecules and assaying for binding of the nucleic acid molecule with components within the test sample. Such assays will typically involve arrays comprising many genes, at least one of which is a gene of the present invention.

Conditions for incubating a nucleic acid molecule with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid molecule used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or array assay formats can readily be adapted to employ the novel fragments of the Drosophila genome disclosed herein. Examples of such assays can be found in Chard, T, *An Introduction to Radioimmunoassay and Related Techniques,* Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G. R. *et al., Techniques in Immunocytochemistry,* Academic Press, Orlando, FL Vol. 1 (1 982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology,* Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The test samples of the present invention include cells, protein or membrane extracts of cells. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing nucleic acid extracts or of cells are well known in the art and can be readily be adapted in order to obtain a sample that is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the nucleic acid molecules that can bind to a fragment of the Drosophila genome disclosed herein; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound nucleic acid. Preferred kits will include chips that are capable of detecting the expression of 10 or more, or 50 or more, or all of the GPCR genes expressed in Drosophila.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers, strips of plastic, glass or paper, or arraying material such as silica. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the nucleic acid probe, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound probe. One skilled in the art will readily recognize that the previously unidentified GPCR genes of the present invention can be routinely identified using the sequence information disclosed herein can be readily incorporated into one of the established kit formats which are well known in the art, particularly expression arrays.

Uses of vectors and host cells

The recombinant host cells expressing the proteins described herein have a variety of uses. First, the cells are useful for producing a GPCR protein or protein that can be further purified to produce desired amounts of GPCR protein or fragments. Thus, host cells containing expression vectors are useful for protein production.

Host cells are also useful for conducting cell-based assays involving the GPCR protein or GPCR protein fragments, such as those described above as well as other formats known in the art. Thus, a recombinant host cell expressing a native GPCR protein is useful for assaying compounds that stimulate or inhibit GPCR protein function.

Host cells are also useful for identifying GPCR protein mutants in which these functions are affected. If the mutants naturally occur and give rise to a pathology, host cells containing the mutations are useful to assay compounds that have a desired effect on the mutant GPCR protein (for example, stimulating or inhibiting function) which may not be indicated by their effect on the native GPCR protein.

Genetically engineered host cells can be further used to produce non-human transgenic animals. A transgenic animal is preferably a mammal, for example a rodent, such as a rat or mouse, in which one or more of the cells of the animal include a transgene. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal in one or more cell types or tissues of the transgenic

animal. These animals are useful for studying the function of a GPCR protein and identifying and evaluating modulators of GPCR protein activity. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, and amphibians.

A transgenic animal can be produced by introducing nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Any of the GPCR protein nucleotide sequences can be introduced as a transgene into the genome of a non-human animal, such as a mouse.

Any of the regulatory or other sequences useful in expression vectors can form part of the transgenic sequence. This includes intronic sequences and polyadenylation signals, if not already included. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the GPCR protein to particular cells.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of transgenic mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene can further be bred to other transgenic animals carrying other transgenes. A transgenic animal also includes animals in which the entire animal or tissues in the animal have been produced using the homologously recombinant host cells described herein.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al. PNAS 89*:6232-6236 (1992). Another example of a recombinase system is the FLP recombinase system of *S. cerevisiae* (O'Gorman *et al. Science 251*:1351-1355 (1991). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein is required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al. Nature 385*:810-813 (1997) and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0 phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring born of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

Transgenic animals containing recombinant cells that express the proteins described herein are useful to conduct the assays described herein in an *in vivo* context. Accordingly, the various physiological factors that are present *in vivo* and that could effect ligand binding, GPCR protein activation, and signal transduction, may not be evident from *in vitro* cell-free or cell-based assays. Accordingly, it is useful to provide non-human transgenic animals to assay *in vivo* GPCR protein function, including ligand interaction, the effect of specific mutant GPCR proteins on GPCR protein function and ligand interaction, and the effect of chimeric GPCR proteins. It is also possible to assess the effect of null mutations, that is mutations that substantially or completely eliminate one or more GPCR protein functions.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the art of molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Gene	Transcript	Similar to	F	W	V
			E Value		
A. Rhodopsin-like receptor family					
Opsin-like					
CG5638	CT17820	Rh3	e-57	e-14	e-44
Receptors for biogenic amines and related compounds					
5-HT receptor-like					
CG8007	CT24060	5-HT2	e-21	e-35	e-27
Muscarinic Acetylcholine receptor-like					
CG7918	CT23924	mAcR-60C	e-32	e-49	e-41
Unclassifiable biogenic amine receptor-like					
CG17004	CT37739	5-HT7	e-18	e-20	e-28
CG7431	CT22855	CG16766	e-70	e-45	e-37
CG16766	CT37292	CG7431	e-53	e-31	e-22
CG12796	CT38338	CG6919	e-23	e-20	e-30
CG6919	CT21432	CG6989	e-82	e-46	e-57
CG6989	CT21650	CG6919	e-82	e-39	e-46
CG7078	CT21843	CG6919	e-73	e-46	e-38
CG18314	CT41076	CG6919	e-21	e-48	e-25
CG7994	CT24036		>e-10	>e-10	e-10
Purine receptors					
Adenosine receptor-like					
CG9753	CT27563	CG6989	e-19	e-29	e-42
Peptide receptors					
Allatostatin receptor-like					
CG10001	CT28187	EG:121E7.2	e-62	e-35	e-34
FSH/TSH/LH receptor-like					
CG4187	CT13764	CG5042	e-48	e-23	e-24
CG5042	CT16185	CG4187	e-44	e-25	e-23
Gastrin/CCK receptor-like					
CG6857	CT21155	CG6881	e-96	e-33	e-19
CG6881	CT21314	CG6857	e-96	e-23	e-33
Gonadotropin releasing hormone receptor-like					
CG10698	CT29989	GRHR	e-43	e-34	e-39
Growth hormone secretagoge receptor-like					
CG8784	CT25324	CG8795	e-165	e-38	e-35
CG8795	CT25350	CG8784	e-165	e-40	e-37
CG9918	CT27924	CG8795	e-83	e-53	e-33
Tachykinin receptor-like					
CG1147	CT1960	NepYr	e-31	e-43	e-40
CG10626	CT29768	Takr99D	e-53	e-41	e-44
Somatostatin receptor-like					
CG7285	CT22465	CG13702	e-96	e-24	e-51
CG13702	CT33159	CG7285	e-96	e-26	e-46
Vasopressin receptor-like					
CG6111	CT19191	GRHR	e-36	e-26	e-45
Unclassifiable peptide receptors					
CG13575	CT32957	CG10626	e-08	e-08	e-8#
CG14003	CT33559	Takr99D	e-14	e-12	e-14#
CG5911	CT18539	Takr86C	e-18	e-19	e-26#
CG10823	CT18916	CG6857	e-10	e-10	e-12

Gene	Transcript	Similar to	F	W	V
			E Value		
Orphan receptors					
Related to *C. elegans* orphan receptors					
CG2114	CT2366		>e-10	e-27	e-10
CG3171	CT10621	EG:22E5.10	e-50	e-16	e-26
CG5936	CT18637		>e-10	e-18	>e-10
CG8985	CT25824	CG13803	e-162	e-31	>e-10
CG13803	CT33298	CG8985	e-162	e-36	>e-10
CG13229	CT32473	CG13803	e-58	e-35	e-10
Other orphan receptors					
CG9569	CT17758		>e-10	>e-10	>e-10
CG12290	CT19320		>e-10	>e-10	>e-10
CG6986	CT21642	CG16726	e-10	>e-10	>e-10
CG13579	CT32961		>e-10	e-18	>e-10
CG13995	CT33551		>e-10	e-13	e-11
CG7497	CT23019		>e-10	>e-10	>e-10
Calcitonin receptor-like					
CG4395	CT4121	CG17415	e-46	e-16	e-41
CG17415	CT38445	CG4395	e-42	e-20	e-61
CG13758	CT33238	CG8422	e-40#	e-65	e-66
B. Se					
Diuretic hormone receptor-like					
CG8422	CT24513	CG12370	e-122	e-25	e-65
CG12370	CT24959	CG8422	e-122	e-30	e-101
HE6 receptor-like					
CG11318	CT31591	CG15556	e-104	e-11	e-20
CG15556	CT35672	CG11318	e-103	>e-11	e-6
Latrophilin-like					
CG8639	CT8755		>e-10	e-47	e-45
Methuselah-like					
CG4521	CT14539	CG6965	e-30	e-13	
CG17795	CT16507	mth	e-132	>e-10	>e-10
CG6530	CT20339	CG6536	e-104	>e-10	>e-10
CG6536	CT20351	CG6530	e-115	>e-10	
CG6965	CT21585	CG17795	e-12	>e-10	e-10
CG7476	CT22963	mth	e-66	>e-10	>e-10
CG13406	CT32762	CG17084	e-47	>e-10	>e-10
CG17084	CT33414	CG13406	e-47	>e-10	>e-10
CG17061	CT33415	mth	e-80	>e-10	>e-10
CG16992	CT37715	mth	e-32	>e-10	>e-10
C. Metabotropic glutamate receptor family					
GABA-B receptor-like					
CG3022	CT9836	CG6706	e-96	e-56	e-85
CG6706	CT20836	CG3022	e-120	e-61	e-169
CG15274	CT35221	CG3022	e-49	e-156	e-129
Metabotropic glutamate receptor-like					
CG8692	CT5032	Glu-RA	e-152	e-115	e-157
Other					
CG7155	CT22117		>e-10	>e-10	>e-10
CG11923	CT35779	CG17215	e-17	>e-10	e-17
D. Atypical 7 TM proteins					
Frizzled-like					
CG4626	CT9057	fz	e-36	e-54	e-76

Claims

That which is claimed is:

1. An isolated protein consisting of an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198;

(b) the amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said allelic variant is encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 795, 798, 801 (genomic sequences); SEQ ID NOs: 2, 5, 8, . . . , 796, 799, 802 (transcript sequences), under stringent conditions;

(c) the amino acid sequence of an ortholog of an amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 795, 798, 801 (genomic sequences); SEQ ID NOs: 2, 5, 8, . . . , 796, 799, 802 (transcript sequences), under stringent condition;

(d) a fragment of the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein the fragment comprises at least 10 contiguous amino acids.

2. An isolated protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198;

(b) the amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said allelic variant is encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 795, 798, 801 (genomic sequences); SEQ ID NOs: 2, 5, 8, . . . , 796, 799, 802 (transcript sequences), under stringent conditions;

(c) the amino acid sequence of an ortholog of an amino acid sequence shown in the SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 795, 798, 801 (genomic sequences); SEQ ID NOs: 2, 5, 8, . . . , 796, 799, 802 (transcript sequences), under stringent condition;

(d) a fragment of the amino acid sequence shown in the SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein the fragment comprises at least 10 contiguous amino acids.

3. An isolated antibody that selectively binds to a protein of claim 1.

4. An isolated nucleic acid molecule consisting of a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198;

(b) a nucleotide sequence that encodes of an allelic variant of the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said nucleic acid molecule hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)), under stringent conditions;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said nucleic acid molecule hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)), under stringent condition;

(d) a nucleotide sequence that encodes a fragment of the amino acid sequence shown in the SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said the fragment comprises at least 10 contiguous amino acids; and

(e) A nucleic acid molecule that is the complement of a nucleic acid molecule of (a)-(d).

5. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198;

(b) a nucleotide sequence that encodes of an allelic variant of the amino acid sequence shown in SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said nucleic acid molecule hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196

(genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)), under stringent conditions;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in the SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said nucleic acid molecule hybridizes to the nucleic acid molecule shown in SEQ ID NOs: 1, 4, 7, . . . , 190, 193, 196 (genomic sequences) or SEQ ID NOs: 2, 5, 8, . . . , 191, 194, 197 (transcript sequences)), under stringent condition;

(d) a nucleotide sequence that encodes a fragment of the amino acid sequence shown in the SEQ ID NOs: 3, 6, 9, . . . , 192, 195, 198, wherein said the fragment comprises at least 10 contiguous amino acids; and

(f) A nucleic acid molecule that is the complement of a nucleic acid molecule of (a)-(d).

6. A nucleic acid vector comprising the nucleic acid sequences of claim 4.

7. A nucleic acid vector comprising the nucleic acid sequences of claim 5.

8. A host cell containing the vector of claim 6.

9. A host cell containing the vector of claim 7.

10. A method for producing any of the proteins of claim 1 comprising introducing a nucleotide sequence encoding any of the protein sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the proteins are expressed from the nucleic acid.

11. A method for producing any of the proteins of claim 2 comprising introducing a nucleotide sequence encoding any of the protein sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the proteins are expressed from the nucleic acid.

12. A method for detecting the presence of any of the proteins of claim 1 in a sample, said method comprising contacting said sample with an agent that specifically allows detection of the presence of the protein in the sample and then detecting the presence of the protein.

13. A method for detecting the presence of any of the proteins of claim 2 in a sample, said method comprising contacting said sample with an agent that specifically allows detection of the presence of the protein in the sample and then detecting the presence of the protein.

14. A kit comprising reagents used for the method of claim 12, wherein the reagents comprise an agent that specifically binds to said protein.

15. A method for detecting the presence of a nucleic acid sequence of claim 4 in a sample, the method comprising contacting the sample with an oligonucleotide that hybridizes to the nucleic acid sequences under stringent conditions and determining whether the oligonucleotide binds to the nucleic acid sequence in the sample.

16. A method for detecting the presence of a nucleic acid sequence of claim 5 in a sample, the method comprising contacting the sample with an oligonucleotide that hybridizes to the nucleic acid sequences under stringent conditions and determining whether the oligonucleotide binds to the nucleic acid sequence in the sample.

17. A kit comprising reagents used for the method of claim 15, wherein the reagents comprise a compound that hybridizes under stringent conditions to any of the nucleic acid molecules.

18. A kit comprising reagents used for the method of claim 16, wherein the reagents comprise a compound that hybridizes under stringent conditions to any of the nucleic acid molecules.

19. A method for identifying an agent that binds to any of the proteins of claim 1, said method comprising contacting the protein with an agent and assaying the contacted mixture to determine whether a complex is formed with the agent bound to the protein.

20. A method for identifying an agent that binds to any of the proteins of claim 2, said method comprising contacting the protein with an agent and assaying the contacted mixture to determine whether a complex is formed with the agent bound to the protein.

SEQUENCE LISTING

<110> Cravchik, Anibal

<120> ISOLATED G-PROTEIN COUPLED RECEPTORS,
NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
THEREOF AS INSECTICIDAL TARGETS

<130> CL000733PCT

<150> 60/191,638
<151> 2000-03-23

<160> 198

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 5236
<212> DNA
<213> Drosophila

<400> 1
caaaaactaa aacaacaaaa agccggcaag caatggtata caaaataaat aaaaaaaaat 60
atgacaagag aaaatccaca aaaaaatgaa ataaaattga atgaaaaatt taccaaaaac 120
caagttgttt ttattccaag tgtatgtagt tggcaatcaa acgattcaac tgcagttaac 180
atgatttatt cgctacgtaa gtgtgcttgg ctaagttgaa taaatatcat taatatctgt 240
cagctacatt aaaggttcaa acattacata tatatatatc tatatattta tcttaacatt 300
attaaatgtt taagggctta aaaaaaaata tccttaggta cacgggtacg tatgtattat 360
tctgcgtctc tttctatatc ctctcaggat ttgaatatat tttgtattat aaaatattac 420
gcataaaatc atcgggggga acctgttaat tcttggcatt tgattcccgc attttctctg 480
gatatggcat agatattttt ctagggttag ctgtgtgaaa cctaaatatt tttaagttaa 540
atttaaacaa cgatttatac ggctgacaca ttttaaattt ttctgggttc aagcttagat 600
tttgtttaca tgtttttttc ttaatgctta acactatgct ttatatactt tttcaaaata 660
tattttaaaa cttgttcagt aaaaagtttt aaagtgataa ttcttaacag agttttttta 720
gaaacttttt ttttagaaag atcaagcata caattatttg tttttatcaa taattcatat 780
gactagaacc tttcttttct aaatgcattt cgggcatagt atcccaggat aaaaaatata 840
acaggaacat tttaaataat tgaaatcagt atgcatttta acaaaattaa gttataaaaa 900
aaaatatttt gctatcatga aataagtgtg ccggactggt cttcaaaatc tttgctcggc 960
gatgcggtca atgttacgat cgaccctcgc gcggtacgcc ttaaacgccc aactctcggc 1020
tgggcgtggt cgtggccatc tgtccagtat tactcccggc cggcggcggg gtcactaaac 1080
tgggcagcaa cttgttcagc agcgtgcgaa tggcaaaggt gacatcgtga ttggcgaagc 1140
agaacaggaa ggagaccaca aagccctgca ggctgaccaa caccaccgat agcatctgat 1200
agaagtggtc cagctgggtg ccggcatccg gacggtaggg cagcaggaag tgttgcagtc 1260
cgaacagcgg taccagaata atggtggccc gaaccgcctt ccggatggcc aacggtgccg 1320
gctgggcgga ctggggatgt aactttcgca cgatcactcg cagtacgttg atcagaaaaa 1380
ctgttgggta aaatgcatat agctgagcaa tattgtactt ttaaagaaaa ttatgcatat 1440
tactaaagct ggccagcagc gaaagagtta tgggcaccga gaagatccac aagtagagac 1500
tgtcagttat ccagcagctg caataatgtg aaaagaaatg ttaccatgga tatagcgtag 1560
aatgtaatct cttaataaat gcaatgcaaa aaccaaatca atcaggatta tgtcttgtgt 1620
aaaattgtga aaataactta ctgcttattg tccggactac tgaaatgcct ggccagaccg 1680
tagacgatgg ctatgggtat gggtgaaaac cagctgatga caataaacca gcgcatcaca 1740
atcgtgtcct tgacaaagac ctggcgaatg gaaagaagaa ggggcgtcag gatctgaacg 1800
gacattcctg acatctcaac ttgacgcaga cgtaaagcag ttggccaacg ttgcgatttg 1860
ctctcagaca aattgctggg ggtgaaccat tgccatgtca gcggctgcag gaaatcacat 1920
ctgccatgca cagaagtcgc ccataaagaa tgcagacttg tgcaggcaat tgccagattg 1980
ccagacgagt aaggataact caccacaacg agcaccaggt gcaggtgcag gccctcgcaa 2040
aacatccaga aatagttgac caacataaag tagtgcacca ccagatgcag cccaatgcac 2100
cacagctgaa agttagttat gtttatggtt ttgtataatt gcctagctac ttgtagggct 2160

```
atgttatttt catagttact tttctcagtg catcttaccg gattttcggc tatggtttcg 2220
gaacgctcca ccaccagacg ataccataat atccaggcca cacaggtgca ggccagcgag 2280
gcgaacagat gtacatggat gcggatgcgg gtgcagcgca gggatctgca aacggaatgg 2340
aacggaacgg aaagaaacga aaagggatat ccggccgtca gttaccggag ggaaaggaac 2400
cgtgtgggtc gtagacaaac acttacttga agcccaagaa aataataatc gatatgagca 2460
gggccagcag ggagagggca taacccttca cgtacagctc gttgatgaac tggcgaaact 2520
gcagtggcaa ctttcttgat taatgcccat gtaccacccg cttgatgccg ccgaagcact 2580
tacctccagg tcctcatagt ccacgcagtt ggtgtaattg gaccaggtct gattgctaac 2640
cggatgacgg taccaggagc cgttctccag gcagctatga atgcgaaatg cttacgataa 2700
ttgtgactgg ctaagtgcgg actgatgggt aaagggggta tttctaacaa agcgcctgac 2760
aatgcacaag ttacccaact ttaccaggag ccagctgagc tcagttggcg tgtaaagtga 2820
gatgaacgct atccattaaa ttggagatac atacatgtag tgcaagtttt agagtgctgc 2880
tacagtacac ttagggagat tatctctgca agagttgaag aaggttttat ttttctttca 2940
ttttccatca tctttttctg agatatgctg cgagtttatg taatagacac gaatgtaaac 3000
tctataggat tttccctgtt tgtagcgcac tcatacataa acattgtatt ttatgacaca 3060
cagaacacat taaagcaaat gccgaatatg gcaagatatt ttcgaattta tgctcaaacg 3120
aatattacat attgcagagc ttatttgacc atgcgaaatc ctgaacggaa tgaaaatcta 3180
tcctattcgt ggtggcgtaa tatacaggca tattgggcat ttaactcacg tcttgtgggc 3240
cagaaatttt ctgttgaatc cttcgacgaa atcgggacaa tattgactca gcacagtgcc 3300
ggccggagtc cttggccagc agagatatcc gtcaaagttc agggggcaaa agagtcggtt 3360
ctccttctga tttgtggcca tcgtggcact gacgacggtg gccatcagct cctccttctc 3420
cgccagctcc tcgtccacgt ttcgcatgtc aataacgccc tcgttcatcg cggcattctc 3480
cagatatcgc ggcactgggc taaagttctc cggcaaatcg gaatcgtagt ctgtgcacgg 3540
cgacaagtgc cattaaaatg ggtttacttt ttaatgtata aaaatatagc atttttaata 3600
atagtatcta tgggctgcat aactgaaagt gatgtaattc agttttcaag ttcaattcaa 3660
ttgcatgctt gctttacata tgtaaaataa agtggtattt caaatgtaca taacttgctt 3720
aagatagtta atagatcatc gttctttata gcagggtata attcataaat tctggttgaa 3780
aatatatgta catttggtaa aggtgttgag tgctatttaa atttataaat ttgtcctgag 3840
atgccaatga tttgcattga aacgaaaacc catgtttaaa tttgcatccc caagaagtaa 3900
gtcaatgatt ttgtatatat ttaaatacat tactttgaaa tgcattaaat gaaatcagat 3960
gggctgtaat ttaatttaat tctcttctct ttcttatctg tcttatgcct ctgctatttt 4020
ttcgcggcaa atcacgtata cgcctgttgg ccgccactca cctgttgctg ggccaagtga 4080
aaatgatggg tcgtccgtat cgtatgtgac attctggtat cggaacacgc actccgcata 4140
caaatgcttt aaaaatatgc gcagattatc ctgcgtttgg gcctgcgatg ccacatccac 4200
atttgcgac tccgaatctg aagaggttgt cgccattgtg ggttattcaa cgtgactaac 4260
tcttgtcaac ggctaatgca ctaaaaaaaa atataataat ataaatataa ttcaaagaat 4320
tacgggataa ataggataat caactgcaat aatttatcag aggcaatttg aaaggtttta 4380
aacaagtgtt gcttacccat agaaacttgt atgagctagt ttttaaaatt tattctgatt 4440
taaatatatt ttcttatatt taacattagt tctcactgta tttcgccata gcgaagcgaa 4500
cttgagtccc ttgtggccgt ccattcgtgt gtgcatttct aatttggtta ggcttcaaag 4560
agttgcctga gttgggttga tatttgttgt ggggtcctgt ctgcggggaa atggtgccaa 4620
aggaaatgga atatccacac ggggcacaca cacacactca aaatatcgaa gcacctgagt 4680
gtcgtgtgag ccgagtcacc tactcgtctc tcatatgttg taaatcagtt tattgtctaa 4740
ttaatggcgg tgcttaactg tgtcgctatt tctgctaccg tggttgggcg aacttaattg 4800
catttcgtcg cacagagaga caaagcaaat gccacgggtt aaaaactcaa atgaaatgca 4860
ctttgatcgt gggggcacac aaacacacaa acacataaac acactcacac acccgcaccg 4920
aacacgtggc ccaaaaggcg ggatgagctg caaagttctt gcctaacttg gttatatttt 4980
taatcaaaac tcatttcact tgattgagtt tcgacagaag cctgcaccag gcactggaaa 5040
atctagcatg cacgatacat acaaagtaag atgattttcc cgctcccaag ctaaacttta 5100
atttgcggca ttaatttatt acaccgaact gggacagttt cccatgatgg atatgacgcc 5160
tggctaaaca ttcttgcatg ttatatatat ttttgcaagt gtctcggcca ggcaattagc 5220
cgcaggtaag cggagc                                                 5236
```

<210> 2
<211> 1686
<212> DNA
<213> Drosophila

<400> 2
atggcgacaa cctcttcaga ttcggagtcg caaaatgtgg atgtggcatc gcaggcccaa 60
acgcaggata atctgcgcat atttttaaag catttgtatg cggagtgcgt gttccgatac 120
cagaatgtca catacgatac ggacgaccca tcattttcac ttggcccagc aacagactac 180
gattccgatt tgccggagaa ctttagccca gtgccgcgat atctggagaa tgccgcgatg 240
aacgagggcg ttattgacat gcgaaacgtg gacgaggagc tggcggagaa ggaggagctg 300
atggccaccg tcgtcagtgc cacgatggcc acaaatcaga aggagaaccg actcttttgc 360
cccctgaact ttgacggata tctctgctgg ccaaggactc cggccggcac tgtgctgagt 420
caatattgtc ccgatttcgt cgaaggattc aacagaaaat ttctggccca caagacctgc 480
ctggagaacg gctcctggta ccgtcatccg gttagcaatc agacctggtc caattacacc 540
aactgcgtgg actatgagga cctggagttt cgccagttca tcaacgagct gtacgtgaag 600
ggttatgccc tctccctgct ggccctgctc atatcgatta ttattttctt gggcttcaaa 660
tccctgcgct gcacccgcat ccgcatccat gtacatctgt tcgcctcgct ggcctgcacc 720
tgtgtggcct ggatattatg gtatcgtctg gtggtggagc gttccgaaac catagccgaa 780
aatccggtaa gatgcactga gaaaagtaac tatgaaaata acatagccct acaagtagct 840
aggcaattat acaaaaccat aaacataact aactttcagc tgtggtgcat tgggctgcat 900
ctggtggtgc actactttat gttggtcaac tatttctgga tgttttgcga gggcctgcac 960
ctgcacctgg tgctcgttgt ggtgagttat ccttactcgt ctggcaatct ggcaattgcc 1020
tgcacaagtc tgcattcttt atgggcgact tctgtgcatg gcagatgtga tttcctgcag 1080
ccgctgacat ggcaatggtt cacccccagc aatttgtctg agagcaaatc gcaacgttgg 1140
ccaactgctt tacgtctgcg tcaagttgag atgtcaggaa tgtccgttca gatcctgacg 1200
ccccttcttc tttccattcg ccaggtcttt gtcaaggaca cgattgtgat gcgctggttt 1260
attgtcatca gctggttttc acccataccc atagccatcg tctacgtttt tctgatcaac 1320
gtactgcgag tgatcgtgcg aaagttacat ccccagtccg cccagccggc accgttggcc 1380
atccggaagg cggttcgggc caccattatt ctggtaccgc tgttcggact gcaacacttc 1440
ctgctgccct accgtccgga tgccggcacc cagctggacc acttctatca gatgctatcg 1500
gtggtgttgg tcagcctgca gggctttgtg gtctccttcc tgttctgctt cgccaatcac 1560
gatgtcacct ttgccattcg cacgctgctg aacaagttgc tgcccagttt agtgaccccg 1620
ccgccggccg ggagtaatac tggacagatg gccacgacca cgcccagccg agagttgggc 1680
gtttaa 1686

<210> 3
<211> 561
<212> PRT
<213> Drosophila

<400> 3
Met Ala Thr Thr Ser Ser Asp Ser Glu Ser Gln Asn Val Asp Val Ala
1 5 10 15
Ser Gln Ala Gln Thr Gln Asp Asn Leu Arg Ile Phe Leu Lys His Leu
20 25 30
Tyr Ala Glu Cys Val Phe Arg Tyr Gln Asn Val Thr Tyr Asp Thr Asp
35 40 45
Asp Pro Ser Phe Ser Leu Gly Pro Ala Thr Asp Tyr Asp Ser Asp Leu
50 55 60
Pro Glu Asn Phe Ser Pro Val Pro Arg Tyr Leu Glu Asn Ala Ala Met
65 70 75 80
Asn Glu Gly Val Ile Asp Met Arg Asn Val Asp Glu Glu Leu Ala Glu
85 90 95
Lys Glu Glu Leu Met Ala Thr Val Val Ser Ala Thr Met Ala Thr Asn
100 105 110
Gln Lys Glu Asn Arg Leu Phe Cys Pro Leu Asn Phe Asp Gly Tyr Leu
115 120 125
Cys Trp Pro Arg Thr Pro Ala Gly Thr Val Leu Ser Gln Tyr Cys Pro
130 135 140
Asp Phe Val Glu Gly Phe Asn Arg Lys Phe Leu Ala His Lys Thr Cys
145 150 155 160
Leu Glu Asn Gly Ser Trp Tyr Arg His Pro Val Ser Asn Gln Thr Trp
165 170 175
Ser Asn Tyr Thr Asn Cys Val Asp Tyr Glu Asp Leu Glu Phe Arg Gln

```
            180                 185                 190
Phe Ile Asn Glu Leu Tyr Val Lys Gly Tyr Ala Leu Ser Leu Leu Ala
        195                 200                 205
Leu Leu Ile Ser Ile Ile Ile Phe Leu Gly Phe Lys Ser Leu Arg Cys
    210                 215                 220
Thr Arg Ile Arg Ile His Val His Leu Phe Ala Ser Leu Ala Cys Thr
225                 230                 235                 240
Cys Val Ala Trp Ile Leu Trp Tyr Arg Leu Val Val Glu Arg Ser Glu
                245                 250                 255
Thr Ile Ala Glu Asn Pro Val Arg Cys Thr Glu Lys Ser Asn Tyr Glu
            260                 265                 270
Asn Asn Ile Ala Leu Gln Val Ala Arg Gln Leu Tyr Lys Thr Ile Asn
        275                 280                 285
Ile Thr Asn Phe Gln Leu Trp Cys Ile Gly Leu His Leu Val Val His
    290                 295                 300
Tyr Phe Met Leu Val Asn Tyr Phe Trp Met Phe Cys Glu Gly Leu His
305                 310                 315                 320
Leu His Leu Val Leu Val Val Val Ser Tyr Pro Tyr Ser Ser Gly Asn
                325                 330                 335
Leu Ala Ile Ala Cys Thr Ser Leu His Ser Leu Trp Ala Thr Ser Val
            340                 345                 350
His Gly Arg Cys Asp Phe Leu Gln Pro Leu Thr Trp Gln Trp Phe Thr
        355                 360                 365
Pro Ser Asn Leu Ser Glu Ser Lys Ser Gln Arg Trp Pro Thr Ala Leu
    370                 375                 380
Arg Leu Arg Gln Val Glu Met Ser Gly Met Ser Val Gln Ile Leu Thr
385                 390                 395                 400
Pro Leu Leu Leu Ser Ile Arg Gln Val Phe Val Lys Asp Thr Ile Val
                405                 410                 415
Met Arg Trp Phe Ile Val Ile Ser Trp Phe Ser Pro Ile Pro Ile Ala
            420                 425                 430
Ile Val Tyr Val Phe Leu Ile Asn Val Leu Arg Val Ile Val Arg Lys
        435                 440                 445
Leu His Pro Gln Ser Ala Gln Pro Ala Pro Leu Ala Ile Arg Lys Ala
    450                 455                 460
Val Arg Ala Thr Ile Ile Leu Val Pro Leu Phe Gly Leu Gln His Phe
465                 470                 475                 480
Leu Leu Pro Tyr Arg Pro Asp Ala Gly Thr Gln Leu Asp His Phe Tyr
                485                 490                 495
Gln Met Leu Ser Val Val Leu Val Ser Leu Gln Gly Phe Val Val Ser
            500                 505                 510
Phe Leu Phe Cys Phe Ala Asn His Asp Val Thr Phe Ala Ile Arg Thr
        515                 520                 525
Leu Leu Asn Lys Leu Leu Pro Ser Leu Val Thr Pro Pro Pro Ala Gly
    530                 535                 540
Ser Asn Thr Gly Gln Met Ala Thr Thr Thr Pro Ser Arg Glu Leu Gly
545                 550                 555                 560
Val
```

<210> 4
<211> 15572
<212> DNA
<213> Drosophila

<400> 4
gccgtaaact atggcgatgc cccactgtgc ccgcacacaa attccgccac ctcgctccga 60
aatcgtgtgt atttggttaa taataatcta gttgggcatc gattgttcgg ccaaaatgaa 120
gccactatcc gtggcggagt gttaattaat ttgcacttaa gcgcagctgc tggaagcact 180

tgatagaagt gggcatgccg gttctcttgg gaagttggct tactatagtt tagcgttaag 240
gcgttaactt gtggccttta aagttcttag agaagacttt ctgcagctgc tcaggctgtt 300
cattgcggca tctagcataa aatttagagt ttgaaagtaa gtttctatga aaaaaaaata 360
cacgtatata taagtataac aagaaaagga ccaaatcaaa gtttatatac cattgctgta 420
cttatatatt tgaatattaa atatacacaa ataaatatat accgaataag aacactttat 480
gattaaaagc aacaattatt gattaatcta cttctaccta aactggtgtt gaaactggta 540
taatgaaaaa caactaaggt acatttactt ccatttaatt ttctattggc agctatataa 600
taaacagttc cgatccaacg tgaaccgacc tatcaccatt tttgataaag ttttacgcag 660
aaacgtaaaa atatctaaaa gcatgctatt ttgtatttgt atttttgtat tttgtagaga 720
aaagttgtag atctatattt aattttgcac acatgagcat tttaatcagt tttgtataca 780
ttatttcgaa ctcttcacca acagaataaa atgaaattaa aatggtccga tgtcgggggc 840
caaacaaaat aaaggcgtat tgagttgtgc catgtttcag gtggaaaaac cagctgcaat 900
ataattgctt ttgaataatg cgaatttcgg ggggaagttt cttttaccag cacaatataa 960
acaaaactgc tcaactattt ctcgatattt acattttcag ggtccatcag caacatagat 1020
gacgcggagt atcactgcaa taagacgcag gtgcgcaagg tcatctcggg ggtggtgggt 1080
gccgcctcct cggtcacatc catacaagtt gccaatttat tgcggttatt ccgaattccc 1140
caggtgagtt tccattgtgt gatcactaac tgttcgactc agtgcgtgcg aagtccaaag 1200
gtcatggcat aaaagggcag tcctttttgt ttttagttca gattttgctt tatgctcgca 1260
actacatgta tatgtgctag tttggccaac aaaagtttct cctatttcgt aattaaaatt 1320
tccaggagcc acatggtaac tgttgcaatg gcttgcaggt ctttcggttg gctttgccga 1380
ttcctaagtg tctgcacgtg ccactaaatt agtcgagtcc cgcaggccaa ctagaggaac 1440
atagtgctgc agttgccaag aagtgtacag tagggactct cagcaatcat aacaattaat 1500
gtaattattt tactatatgt tcaatttatc tttaatatac aattccaaaa tgatgattct 1560
tagctgctaa agaactgact tacatattaa tttcctacca aaaccaataa caaactggac 1620
agatcgcttc agttgagtac ttcgactatc agatacccgt cacttgcatt accgcttaaa 1680
aagcttgtat ctagcattgc tcgtatttca actgcaaacg acattgcagt tttttgggtg 1740
ttcctcatgt agtgactaga gcgtgaaaaa cgccacaaac taatatccat ttctgattca 1800
ggaaggaata taacattttc cgagtgccgg catgaatcac gtgcaggaag ttgaaggaga 1860
ttgagtgccc ggttcggctg ctccagctac gtgagctgga gccaaaagtc aacaacgcga 1920
caataactaa gtagccaact gccagatccc ttctcgatat atatatatag tatgtactca 1980
ctactacttg gatgccatta aacgcacacg aagtgccagt cgaaggagca gcagccgttc 2040
catgcaggcc cagctatcgt cccgaaagga cctcataaat cagcaataaa catttgattg 2100
catttggttt gcacgcgtgg attctgcgct ggcttcgcat aaacagcact ttcgagcaaa 2160
ttttgtgtgt tctgcaattc aaatttctgg tcgatagcga cttagtgagt ttcgtttagg 2220
cttagcggaa attggctgca atgcaaaatg ttcgcaagga cacaattcaa atgacaatgc 2280
agttaaaaat tccgtaaatg caatatttta tggcttatcg tatttggata acaatgtttt 2340
aaatgcttga acaaagcatt gctgtctttt ttgtttaaag taaaattttt gccgtcagaa 2400
gtttgtgcat tgaattaaat ggttattcaa acattacatt ttatttatct tgaaaaaagt 2460
acaaaaatat gtattaaact cgtagtatta tatataaatt aaatatattt tattattttt 2520
tgtttttatg gactgagtaa tagcacttga attttccaac tagtttttca caatttttcg 2580
tactttaata aaagtaaata tttgcatctc tagtgtcaat tatcaaaagt caacaaagcc 2640
aaagctacac aataatttg ctataactcg acgctttaaa cgctcatgaa aatcaaaagt 2700
aattctgatt aaataaaagt gagaatgcca taaaccatag gtaaattaga ctcagaaata 2760
aggtctggta tgagttgcaa tgcctgtctt gccatgagca tttgacaagt aatcaagcac 2820
ttacgcagaa gctgttgact aatggtaatg acacttaatt agaagccaag tagttcatta 2880
aacacttgca gttgacttgc cacgccatcg acgatttgcc ccaggatgca cctacttcct 2940
gccccaagtg tttgtgggcg gatccttctc taaccaaacg gtttatgtgt gcgtgaaatt 3000
ggtttcccgg cttgggccag gtcaccacag ttggcctttt gctgcagttg ccacagcttt 3060
gggtttcgca aggtaatgcc agctgatgtt ttcttcgaca gcagcattgt ttgtggtttt 3120
tggaaaacag cttaagcgag aaaatatgag cgaaaaactg ggaaggagct tacgaaaggt 3180
tctgcagaag cgatagccca gaacgagtgc gaaaaaatat gaaaattaaa tggcgacctt 3240
aaagtgaaag ggttcttac ataaaatatt gtagcacact tgccactttg tgccacccaa 3300
gcaatagcgt ttccaactaa aatgttatcc catcatattg taacccgcct caccacattt 3360
tttccatcca ggtgtcgtac ttctccacca gccccgagct gagcaacaaa cagcgattcg 3420
agtacttctc ccgcactatt ccctcggatc actaccaggt gaaggccatg gtggagatag 3480
tgaagcggat gggatggagc tacgtgtcga ttatctacga ggagagcaat tacggaatca 3540
aggtgagtgt ttatgtatgg aactttgatt aaaatggaag aaaggagtgc atctatggct 3600
gagaaatgga gttaaagaat gagaagctct atgttttaaa agagcttgtt aaaattcaaa 3660
tttaatgtat gttgcttcat tttactttgg atacttcttg gataaaatct atttggata 3720
accattaaac attgaaaagt caataataag taagttcaat aagtttcagc cacttaaatt 3780

tcctcttttt caaaggacta cataaatttc acctaatagg gcaaatcatt taactagtaa 3840
tccttaaacg gtggaacttt aaaataatgc caagtgccgt ttaaaagcgt tagcagcact 3900
ccattatcat atgtttatat ccttagccgc ttgcagtgta aacgatatac attttaagtg 3960
aacgccttgt atcctatcca acttgtacct gtgtaaacat tttacccatc ccaccacatg 4020
cgtttttatt gctcacccac actgcacttg ttttcaggct ttcgaggaac tggaggagct 4080
cctcgcacgc cataacattt gtattgccat caaggagaag ctggtgaagg attcgggagt 4140
ggccgaggac atcgcctacg acaatatagt gcagaagctg ctgacgaagc cgcgggctcg 4200
aggtatgtaa ctggagagtg ggaagttggg attgcagaga gggatcggag gcgatgtggc 4260
ggaggaaaag tgcatttcca accggaacag cgactgcagt ctgccagtca aaatgcaatt 4320
taaaatgaaa atgtcaaagc gaaggcgcag tgcaaatgcg aaaggaaagc ccttcgcaac 4380
gcccgcacaa tgaatgcgat ccggagtcga gaactatgct aggatttttg gctgggcgac 4440
ggaaattgcg taaattaaat gaaaatattt aaaaactccc tcgccagcgc gccgttgtgc 4500
catctgtgcc agcgggaaat taaaaagttc acgacaaaac aaacggaaaa agttatttgc 4560
aaataaattg agagatagtc gcccagaagt gggcaaactg atacgtccaa tactgtagtt 4620
gtgcaaaaaa gtgtgtattt gttatggcca aaaatgaagg tcctgtctat tgtccgtgga 4680
caaaaagcca ggagtacttg gcgtccctga ggtctgatga atgacaccac ctattcgaca 4740
tggcaagcaa ttaaaaattc cagtcagggt tggccaagtt caaatttatc ggaacgattt 4800
ttcatcctaa atgaaaacga tttttccaca tggcattgcc ccaatctgcg ccaatcatgc 4860
gattcagcag cgagataaac aaaaaataaa tgtcaatcaa tttgactttt cacattttgt 4920
gttcccctgt cgaatatcag ctgccattcg cagacgccga aaagaaatat aaattaaaaa 4980
taattaccgc ccacagcttt gcgcatatca ttagtcagtg cctcgaaagg gtgaagaata 5040
tcagggtcgt gcccgccaat aaaaagcaca cgggtcgacc aataacaatt gacttttgtc 5100
tggcaaacaa tttttgattt acataaatgc gccatcagga gtcgtaaagc agtcggtcgc 5160
ccaagaaaag ggttcctccg ccaccattat ttattgtgat ttgtgatttg acgccgctta 5220
cagggcgcta gccgcctcgc aacgattaat caatttgaca aagtattttc cctggcatca 5280
gttgtttaat tttgcatatc gagggttctt tcactgctcc tggtcctggt cgtggttctg 5340
gtgggggctc cagttgggtc cctgaaaaat gatgcccttc cggatccgta actgcggctg 5400
gtaactgccg ccaggggtga actttacgct tgttcggatc cccacacatg tcactccgga 5460
gctaaatgga atgtgggcgg ggggccgaca tgtttttcct ttttatcgcc aatgtggcaa 5520
gctgatgcga tttcagccag gacatgaccg gcccccaccc gctttagatc actttctggc 5580
ccggctttaa agtttccccc taatctgcat aaatatcctg ccgagtgggc tcttaaatta 5640
aattttatat ccctgctaag aataaccgtg gttctcgact ggtataaaca ggctactgtg 5700
tacgagaaag ccgctcaagt gcatgtctaa ggcaaagtga aaatccaagg aaaatatcct 5760
taacttcaca aaagtatcga caagttttcc cgttgtcgcc gccgcttcct ttgaataaga 5820
ttccctcctc ttacttcgag tggcggagcg gcagtggcaa atcgataccg ggacatgttg 5880
aaggggcaaa aaattatggc taagacaaag gacctcttgg gcggaccgtt aagaagtgac 5940
aactgcagct cacgtatccc atacaagttt gtccagagcg aaacaataag ggctgaaagc 6000
gttggaaaat ttggtaagca cctaaccaaa ggatagtttt agaagctcct gatttccatc 6060
cgtagaacac aagcgactta ggtgtcggct ttaccaattg gccattggga acgttgtttg 6120
cttggagttt ctaatgatat cttcacactt tttgtggact ttctgttcgc ttcttgggac 6180
caattaaaat attgttttcc ttgcaaattg aagaacaata ccttggcaga aaatctaaat 6240
atagtaaaga tacgtcgcca aatggcgaac atctgaaaaa gtggttgcat tacatcaata 6300
tatacccttc gtcatgtgta ccccttgtac cccccatacc cttgttttat agacttgttt 6360
tacgcctaac ggctatagta atccaactta atttatatag catttgcgta attaatattc 6420
aaagcgaaat aatgggctgt catttaatac tataaaaaca ctttaatccg ccctcggaaa 6480
ggagaacaca agagcagttg gcagctgcag attgttttct tttgcgcaca aaagccgctc 6540
attcactatt gattgcaatt aaaattaaat tctttcttcc tcctttcgca tccttttaat 6600
acgccgctgc ccagggggcca ttatcttcgg ttcggatcaa gaggtgcgcc aagtgatgcg 6660
ggcggtgcgc cgggcgaatg caactggttc cttctcctgg atcggatccg atggctggag 6720
tgcccggaat ttggtgtccg acgactacga gccggaggta ggtgaactat tccgtcagac 6780
cggatgacgg gccagttacc gtcagccgga aattatgagc ctaaatgggc ggaacaactc 6840
ggtttgatta atgactcgga tttcggtcac atactcacgc acgcacatac gagggcccaa 6900
ccccattaca tgataaccaa aaaattataa atgatggcgg ccgaaagaac actttcacac 6960
gccctttccc atgatgaatt atggttcaat ttggattaat cgtgcctttg tctgtgcgtg 7020
gctaattaat ttgtctaagc ccttcaagcc cagaaatggt cctcctctct gatcaagctc 7080
attaagtata cgccgcgtgt gtcttggtgt catcaacagg tggaaggcac gctgtctgtg 7140
cagccgcagg cgaatccggt gcgcgggttc gaggagtact tcctcagtct gacggtggag 7200
aacaatcaac gcaatccctg gtttgtgggt aagtgaatta ccgattaaaa ttaagctggt 7260
gttcattaat gctctgatat aagtagtcct ttattagtat ccccttttcgc taaatggtta 7320
tcttgatttc agaattctgg gaggatcact ttcagtgccg atatcccggc agtacgagca 7380

ctccgtacaa taactacacc aaacagtgca ccacaaagga gcgactttca cgccagaaca 7440
ctgacttcga ggaccagttg cagtttgtca gcgatgcggt gatggcattt gcatacgccc 7500
taaggtacga ttaagcttat tttggaagca acccactatc attttaaacg tattaagaaa 7560
ataatagttc ttttgatcaa ccctttttagt ttttatcaac ctatggagta gcaataaata 7620
aaatgtaatg tgatttcctc tccttaggga tatgcaccgc gacctgtgcg gcggaggtcc 7680
ttcgctctgc gaggccatga agccaacgaa gggtgcagac ttgctgaaat atttgcgaaa 7740
agtggagttt gagggtgagt gagtaataga agccgtgtat ccacttttgc cggataaaag 7800
caagttaaat taacggagta actattcccg ggaattttca aaacaggatg tccgactgac 7860
tcagatccct ttgctctcgc ttagttcctt cagtcggcaa gtgatgcgca catagcttaa 7920
gtggaaattt cctgccaacc cgtccttgca gttgggccaa aagccccgca cacacactta 7980
cacacacaat ggcatatctc agcctgcgag agtcctacaa ggcgcattgg ccatgctttc 8040
gagtggaaaa ttgacttgct tgtcatcgtc attgaactcg attgaaattc aataaattac 8100
acagacagac ggacgggaag gagccaaagg ccctgactga ctagcgcaga ttgtcaattt 8160
tctgataggg cgagcgggca agcggcaaag tcggcgaagt cgtctctggc aggcgtgaaa 8220
ccataaaccc atttaactcc tattcgatca atccctccgc aggcctcagc ggcgacgagt 8280
tccgcttcga cgggaacggc gacggtcccg cccggtacaa catcatccac ttcaagcagt 8340
cccaggcggg acagtaccac tgggtcaagg tgggcgagta caccgagggc gagctgcggc 8400
ttaatatgac gggtaagata aagcccccag taggaaatac ttaaaagtat aataccagga 8460
acccttcaac attaagttta cctgactgaa ttcgtgcatc cttagtctcc tgtaacctaa 8520
tgtattcgtg ctcactgctc ctcatcaata ctaattttga tctcattatc ccgctaaccc 8580
agaggtgaag tttaagcggc tgagtcccaa gccgccggag tctgtctgca gcctgccatg 8640
tcttgtgggt caggccaaga agtacgtgga aggcgagagc tgctgttggc actgcttcaa 8700
ttgcacaacc tatcaagtaa gtgaagccca tctgaggcca gtaatttatt tgttggccct 8760
tcgaatagat tgaaatgtag cgccaagtga attattccgg aactaacgcg ttgacgggag 8820
atgggaaaaa tggggcaaaa acaattaagc catcaggcct gaaaggggct gcaagtgaat 8880
ggcaaacatg ccgcctttgt gctcgcccca tccatttatc aagtgttcct ttgtgcggtg 8940
cgtccattgc agatccgcca cccggatgac gagacccatt gcaagctgtg caagctgggc 9000
actcttcccg acgcccacaa gcagtactgc cgccccatcc cggagatata tttgcgtccg 9060
gagtccgctt gggccatcgg tgctatggcg ttcagtgcga ccgggatctt agtaacgctc 9120
ttcgtaatgg gcgtatttgt caggtgagca gcgcatacac cgtgataaaa acgggtatag 9180
gaaattactt ttaataatat ttaatagtta ttaatggctt ttataaaacc atttgaaaca 9240
attcagtctt aagatcgaac atttaacaat gaaatttgct ttctaaatcc aaaaacccat 9300
caggatgtta tattcctttt ctattaagac catttttaat aatacatagt ccaatattat 9360
cataaaacct tgttaatgta tatattttct caacttcaaa taatttagac aatatataat 9420
actttaacaa taattaaata atttgaagat aaggtaacgg gtatattata gtcaatgcat 9480
tcgttcttgt gtttataagt atagttattg ttatactagt gacaagctaa aggatctaga 9540
tcaagctgat ctgagtgcca aatatttcat gagggcctgt tttctagcta tgtttttctc 9600
tgactgtatt tattggcaca caattgtccg cttgcatgga tatggaaagc aattaaaaag 9660
tttcaccctt tccgggcaaa cttcgccaca aagttaaccg gcagacacat aaaaatgatg 9720
cattaattgg cgcaagaaaa cgtcgtcgca agtttgaaac tgaagtaacg cctttccctc 9780
cttcaccttc cacccaacag gcacaacgac acgcccatag tgcgtgcctc tggtcgggag 9840
ctgagctaca tcctgctggc ggggattttc atgtgctacg gtgtcacctt tgccctggtc 9900
ctgaagccga ccaacatagt gtgtgccatt cagcggtgag tgtgcaaatt gccgagatga 9960
tgcgtgtgaa ttttcctccg aaactccctc ttttccgctt ttccgcttcc acacaggttt 10020
ggtgtcggct tctgcttcac ggtcgtttat gcagcactgt tgaccaaaac gaatcgcatt 10080
gcacggatct tcaaagcggg caaacaatcc gccaagaggc cctccttcat tagtcccaag 10140
tcccagctgg taatctgtgc gtgcctggtt agtgtgcagg tgagtgtgtg gagtggtctg 10200
gtggcgctcc tggcggtgac cgacggggcg gatgtggaat cgagatgccc tgcgatgact 10260
ccgtgtcatt ctttccattc actccattcg cagatactca tcaacggcgt ttggatggta 10320
attgcgccat cacatgccat gcaccattat cccacacgcg aggataatct gctcgtctgc 10380
gactcgtaca tcgacgcctc ctacatgatt gcgttctcct acccaatctt cctgattgtt 10440
atttgcacgg tctacgcggt gctcacgcga aagattccgg aggctttcaa cgaatcgaag 10500
catattggtg agtaattgag ctttttattt ccagaagagc tataattcgg cgtaattggc 10560
attgtattgg aagggtctgt acaaaaagcc atgcatgggt ttgcattgtc tataacctgg 10620
ttcattatag cagcaatgct attaactggt ttatttagtg cataaaatag cacaacatgt 10680
atgaatgaac tccttataat ttagttatgt gttttgtatt tgggtaattt ggttatttcg 10740
aaacaataca cgcatgaaat atttataagt ctatatttta tgcattgcaa cctcagagtt 10800
ctagtgcaga aatgttttat aattatgaag ccataatgcc cataattggc cactttgatt 10860
gattacttaa ccctattgtt acaggattca ccatgtacac gacttgtgtg atttggctgg 10920
cttttgtacc gctctacttc ggaaccgcca accacgttcc gctgcgcatc accagcatgt 10980

```
cggtgaccat cagcttatcc gccagcgtga ccatcgcctg tctcttttcg cccaaggtga 11040
gtgagtccgg atagctggga gtgggttggg cttcggcttg gggggattca tatccacacg 11100
ggtcctctgt aaattgattg gtaattgcgt tatgcattaa ttattgcgca tttatctggc 11160
gaccgttgaa tgatttttgg tgccacagta tgcagaagcc atcggcgatt ttattccatt 11220
tcattccgtt tattcaggac gcgtggacga gtgtttcatt aaaatgcaat aagacggaag 11280
atgaggacag gcggaagcca agggtttggc aattcaacaa ccggctgcat cgattgcgat 11340
gtgttgcttc caaccatatt gtttttaatg cctacccttc ctcttggtcc atattccttg 11400
gtccttgccc atttcatgat gtttccccga gatgaagctt aaatgtcaac aagccatgtt 11460
ctcccacgga gctgaactgg tcatgatctc ctcccaaagg cgcgcgtcct ttatgttggc 11520
atactcggtc cccatgccaa catatttttt aatacgccct gctattagct attaagatgc 11580
aagtaaactg ccagacaagc caaatatagc cagccagttg agttcatgga ctagatgatg 11640
cagcatggca aagaagctgt ttaccccgat attaaacgtc aaggaggaag cttacattta 11700
agcactgata tcttcattac acggaaacac gacctttgac agcaatgtgc aatgtgaaac 11760
gcaggaaagg aattttataa gaaaatgtcg aatagcttgt ttgattattt tttttgtata 11820
agtattccaa gcatgaaagt tagatcagaa aatccaagtt atattggcaa ataatgaaag 11880
tctcttaatt gtatggctat taaaggggaa acacttttct agttagtttc caaaggcttc 11940
catataacac ttaaagatgc cagaaaggac cttcgccagt gtctatgcca ttgaagagca 12000
aggaggtaga aatagtaaat cagccagaaa cttgattagc aacggagcgg agctcccaca 12060
accgcacacg tcatagttct tggtcccgga gcggagcaag ctacagttta tggtctgtct 12120
gcatattccg ggtgcaaagt agaacaaaac ccggggaggc gaagggccgt aaacactcag 12180
caaaaggcta ttattaaatg tcggagtgga gcatacttct tggcccagtg gctgctaacc 12240
gggccaactg tacgctatca gatgtcacct aatgcacggt tggtttgcga cgaggaagga 12300
acactgtact gcgcacacgg caaatacagt gtgcacaagc acacgaaccg caccgtacac 12360
cttgagccat aaaaatgcaa tttacaactg gttgtgtctg gcaaaaggaa aacaataaaa 12420
ggctgccatg aatgcggcat gccacacaac aactggggat caccggacca agaggcaaca 12480
aactaaggcc ctggcgcaaa ggtgccggca gccacgcctc cttccgcagg acaaaggctg 12540
gcaaccgagg tggcaagcca ttccgacttc cgactgccac gcgccctttc tgctgtccgt 12600
ccggctctgc cgcagtttgt tttcgaactt gccgcgtttg tttgcggtga aaacaataac 12660
aaggggcgtc caaatggcag ccaaacaatg tcaaagcaaa gcggaaatga cccaccgtgc 12720
aggctacgta acagtacgca gcgtatcgca agttgaaacc tggagcactg ctcgcaattt 12780
gcgctaaatt ttcactaagc gttttataga tggtgccatt ctagtaagga gtatgtgtgc 12840
tctgaaccag gaactcgctc acgtggccaa aaacaaattc gatttggctt taaatatata 12900
ttagaaagtc aagacacctc aagaaattgg tgggaaagtg ttttcatgta ttgaatagta 12960
atcttattag cggtattttt tacacctcac tagtagagtg aaagggtaga ctagtttcgt 13020
tgaaaagtat gtaacaggta aaaacaagca tttaaagtat aaagattatt gattggaatc 13080
aatagcagat acgatctgac catgtccgtc cattcgtatt aacgtcgaaa ctaataagtt 13140
gagataaagc atactgattc cagagacgta gacgcagcac aagtttccgt attcaatatc 13200
cgttccggca gatctttcta tagtcgggga actcgactat agcgttccct aaattatatt 13260
taaattaatc aaacaaaaat gaacgaatgc gtgaacgtgt tatgatttaa tccgtctatt 13320
acattgttca ttgttatatt attcttaaag tgtaagtatg gtaatatggg gctcaggcgt 13380
gcacttctgc tccctgtaag tggcccaagt cgttggcctt tgctttgacc acattactgc 13440
tggccgtaga agcaggaaat ggctaatgac acttgaattt ggattaggaa tgggacttgg 13500
tggtgggcgt ggcaggtcac gcactgactt ctaatcaaag ccgtaagata agccatcttg 13560
gggaacaacg aaattggata tcgcctccaa atagctatta cacacaatgg gtcgttcgat 13620
taatgcctcg aacgggaaat taacctaatt ggaaagacgc actgcttttc tgactcgacc 13680
ctacgcttgt gtccttgcag ctgtacataa tactcatacg ccccgagcgc aacgtgcgcc 13740
agagtatgat gccgcctcgc tacggcaaca tgcaccgcac cgctggcaca ggaccctcgt 13800
ccatgatggc cgccgccgtg gtgaccgccg ccacctgtgc ccaggaggag aagatccaga 13860
agcatatcac tcccacaaac acaggtcagt aattaaggga tatcggttaa cgagtcccgg 13920
gtcagtaagc gtagcgctca aaatacacac acttcgtgtc gggcaaatgt attattgaca 13980
tatgtaattt gctgttgaat agtgcaatta agaggtataa cagttgcttg cctcgttaat 14040
cattatcgat caattaatca aattattgaa agtgagtcga ttttatctcc tgacaatgta 14100
ttaatcgtat tttgcaatga atttatatca attaaattat attactttat taacccattg 14160
caccagctgt acctatctct cgtattcatt tttaattttt gtggaaatgc ttcaattaat 14220
tgagaaaagt ttatcataaa ttgggtaata tagcaagtct taaaataact caagctgtaa 14280
aagctgaaca gttttatcaa agtaatagtt aggctgacaa tcaaaaggta tatccttctt 14340
aatttaattg attttaaaat gactaggata aagaataaaa ctgaattttt tccattcact 14400
acccataact tgcagaaaac tcgctaaaga cgaagaaatt gtgcgaaatg gccactcaga 14460
caatttccag tataatcacg tcgctggaca tcaatgcgta caatcaaata ccatatgccg 14520
atcaatatgt gccaaacaat gccacaacgg ccacgaccac gcccacagat aagggcagtg 14580
```

gcaacggcaa cagcaacagc aatagcaata cgaatggtag tagcaatagc aacagcaaca 14640
gcaatggcga tggcgatgtg gcaacagagc aagtcgtggc caacaataat aaaatcaatc 14700
agcgacaaca tggccagccg gccgtggcat ttgccatcac atcgactaac aatcacataa 14760
gcggcccagc agcaaccata acaacaacag cggcaactgg cgaaacaaca actgaggcgc 14820
tggtggctac aataacaact ccattggcaa cggtggatgg cagcgaaacc atgtccgtgg 14880
ctcccaacaa tgggcatgga aacaacggga gccaacgtcc gccggtcctg caaacaaact 14940
tatagctaaa cgcgcaacaa ttgcagcaac agctgctgaa ggagcagcag cgacaacaag 15000
ccgagctgca aacgcaacat ccgcagcagc aacatggaac acaaagctct tccagcgaca 15060
ccgccagcac gggcagcggc agtgcggcga agaagcgggc cgcaattccc gtttagctgc 15120
ccccggagat tcgagattct cgcaaccata cgtaagtaga ggaaagtgtc tgttgaccca 15180
agcgattccc ggcatcagtc tagggatgcg gaaaagcaga caagaggagg gcaggaggtc 15240
gtgctaggac gggactaaat tgggagcagc tgggccacag aaatatcttc agatcttaga 15300
aggcgaaata attcatttac cacttttgtt cgcttattaa tttgagtgaa ttctggttat 15360
accttacccc catagacaca aattatctta acgttctttt atataagttt atcgatagct 15420
gtacaacaat gtatctcaat aatgaagata ttcctgtggc acttaaggct ttcttcagat 15480
tttaacatgc acgcaagagt cataagtaag taagagtcgg ctgatcaaaa accttcgcaa 15540
ctcttcccaa aaacaaatgt gtttctccat tt 15572

<210> 5
<211> 2456
<212> DNA
<213> Drosophila

<400> 5
ggtccatcag caacatagat gacgcggagt atcactgcaa taagacgcag gtgcgcaagg 60
tcatctcggg ggtggtgggt gccgcctcct cggtcacatc catacaagtt gccaatttat 120
tgcggttatt ccgaattccc caggtgtcgt acttctccac cagccccgag ctgagcaaca 180
aacagcgatt cgagtacttc tcccgcacta ttccctcgga tcactaccag gtgaaggcca 240
tggtggagat agtgaagcgg atgggatgga gctacgtgtc gattatctac gaggagagca 300
attacggaat caaggctttc gaggaactgg aggagctcct cgcacgccat aacatttgta 360
ttgccatcaa ggagaagctg gtgaaggatt cgggagtggc cgaggacatc gcctacgaca 420
atatagtgca gaagctgctg acgaagccgc gggctcgagg ggccattatc ttcggttcgg 480
atcaagaggt gcgccaagtg atgcgggcgg tgcgccgggc gaatgcaact ggttccttct 540
cctggatcgg atccgatggc tggagtgccc ggaatttggt gtccgacgac tacgagccgg 600
aggtggaagg cacgctgtct gtgcagccgc aggcgaatcc ggtgcgcggg ttcgaggagt 660
acttcctcag tctgacggtg gagaacaatc aacgcaatcc ctggtttgtg gaattctggg 720
aggatcactt tcagtgccga tatcccggca gtacgagcac tccgtacaat aactacacca 780
aacagtgcac cacaaaggag cgactttcac gccagaacac tgacttcgag gaccagttgc 840
agtttgtcag cgatgcggtg atggcatttg catacgccct aagggatatg caccgcgacc 900
tgtgcggcgg aggtccttcg ctctgcgagg ccatgaagcc aacgaagggt gcagacttgc 960
tgaaatattt gcgaaaagtg gagtttgagg gcctcagcgg cgacgagttc cgcttcgacg 1020
ggaacggcga cggtcccgcc cggtacaaca tcatccactt caagcagtcc caggcgggac 1080
agtaccactg ggtcaaggtg ggcgagtaca ccgagggcga gctgcggctt aatatgacgg 1140
aggtgaagtt taagcggctg agtcccaagc cgccggagtc tgtctgcagc ctgccatgtc 1200
ttgtgggtca ggccaagaag tacgtggaag gcgagagctg ctgttggcac tgcttcaatt 1260
gcacaaccta tcaaatccgc cacccggatg acgagaccca ttgcaagctg tgcaagctgg 1320
gcactcttcc cgacgcccac aagcagtact gccgccccat cccggagata tatttgcgtc 1380
cggagtccgc ttgggccatc ggtgctatgg cgttcagtgc gaccgggatc ttagtaacgc 1440
tcttcgtaat gggcgtattt gtcaggcaca acgacacgcc catagtgcgt gcctctggtc 1500
gggagctgag ctacatcctg ctggcgggga ttttcatgtg ctacggtgtc acctttgccc 1560
tggtcctgaa gccgaccaac atagtgtgtg ccattcagcg gtttggtgtc ggcttctgct 1620
tcacggtcgt ttatgcagca ctgttgacca aaacgaatcg cattgcacgg atcttcaaag 1680
cgggcaaaca atccgccaag aggccctcct tcattagtcc caagtcccag ctggtaatct 1740
gtgcgtgcct ggttagtgtg cagatactca tcaacggcgt ttggatggta attgcgccat 1800
cacatgccat gcaccattat cccacacgcg aggataatct gctcgtctgc gactcgtaca 1860
tcgacgcctc ctacatgatt gcgttctcct acccaatctt cctgattgtt atttgcacgg 1920
tctacgcggt gctcacgcga aagattccgg aggctttcaa cgaatcgaag catattggat 1980
tcaccatgta cacgacttgt gtgatttggc tggcttttgt accgctctac ttcggaaccg 2040

ccaaccacgt tccgctgcgc atcaccagca tgtcggtgac catcagctta tccgccagcg 2100
tgaccatcgc ctgtctcttt tcgcccaagc tgtacataat actcatacgc cccgagcgca 2160
acgtgcgcca gagtatgatg ccgcctcgct acggcaacat gcaccgcacc gctggcacag 2220
gaccctcgtc catgatggcc gccgccgtgg tgaccgccgc cacctgtgcc caggaggaga 2280
agatccagaa gcatatcact cccacaaaca cagacgaaga aattgtgcga aatggccact 2340
cagacaattt ccagtataat cacgtcgctg gacatcaatg cgtacaatca aataccatat 2400
gccgatcaat atgtgccaaa caatgccaca acggccacga ccacgcccac agataa 2456

<210> 6
<211> 738
<212> PRT
<213> Drosophila

<400> 6
Met Val Glu Ile Val Lys Arg Met Gly Trp Ser Tyr Val Ser Ile Ile
1 5 10 15
Tyr Glu Glu Ser Asn Tyr Gly Ile Lys Ala Phe Glu Glu Leu Glu Glu
20 25 30
Leu Leu Ala Arg His Asn Ile Cys Ile Ala Ile Lys Glu Lys Leu Val
35 40 45
Lys Asp Ser Gly Val Ala Glu Asp Ile Ala Tyr Asp Asn Ile Val Gln
50 55 60
Lys Leu Leu Thr Lys Pro Arg Ala Arg Gly Ala Ile Ile Phe Gly Ser
65 70 75 80
Asp Gln Glu Val Arg Gln Val Met Arg Ala Val Arg Arg Ala Asn Ala
85 90 95
Thr Gly Ser Phe Ser Trp Ile Gly Ser Asp Gly Trp Ser Ala Arg Asn
100 105 110
Leu Val Ser Asp Asp Tyr Glu Pro Glu Val Glu Gly Thr Leu Ser Val
115 120 125
Gln Pro Gln Ala Asn Pro Val Arg Gly Phe Glu Glu Tyr Phe Leu Ser
130 135 140
Leu Thr Val Glu Asn Asn Gln Arg Asn Pro Trp Phe Val Glu Phe Trp
145 150 155 160
Glu Asp His Phe Gln Cys Arg Tyr Pro Gly Ser Thr Ser Thr Pro Tyr
165 170 175
Asn Asn Tyr Thr Lys Gln Cys Thr Thr Lys Glu Arg Leu Ser Arg Gln
180 185 190
Asn Thr Asp Phe Glu Asp Gln Leu Gln Phe Val Ser Asp Ala Val Met
195 200 205
Ala Phe Ala Tyr Ala Leu Arg Asp Met His Arg Asp Leu Cys Gly Gly
210 215 220
Gly Pro Ser Leu Cys Glu Ala Met Lys Pro Thr Lys Gly Ala Asp Leu
225 230 235 240
Leu Lys Tyr Leu Arg Lys Val Glu Phe Glu Gly Leu Ser Gly Asp Glu
245 250 255
Phe Arg Phe Asp Gly Asn Gly Asp Gly Pro Ala Arg Tyr Asn Ile Ile
260 265 270
His Phe Lys Gln Ser Gln Ala Gly Gln Tyr His Trp Val Lys Val Gly
275 280 285
Glu Tyr Thr Glu Gly Glu Leu Arg Leu Asn Met Thr Glu Val Lys Phe
290 295 300
Lys Arg Leu Ser Pro Lys Pro Pro Glu Ser Val Cys Ser Leu Pro Cys
305 310 315 320
Leu Val Gly Gln Ala Lys Lys Tyr Val Glu Gly Glu Ser Cys Cys Trp
325 330 335
His Cys Phe Asn Cys Thr Thr Tyr Gln Ile Arg His Pro Asp Asp Glu
340 345 350
Thr His Cys Lys Leu Cys Lys Leu Gly Thr Leu Pro Asp Ala His Lys
355 360 365

Gln Tyr Cys Arg Pro Ile Pro Glu Ile Tyr Leu Arg Pro Glu Ser Ala
    370   375   380
Trp Ala Ile Gly Ala Met Ala Phe Ser Ala Thr Gly Ile Leu Val Thr
385   390   395   400
Leu Phe Val Met Gly Val Phe Val Arg His Asn Asp Thr Pro Ile Val
                405   410   415
Arg Ala Ser Gly Arg Glu Leu Ser Tyr Ile Leu Leu Ala Gly Ile Phe
            420   425   430
Met Cys Tyr Gly Val Thr Phe Ala Leu Val Leu Lys Pro Thr Asn Ile
        435   440   445
Val Cys Ala Ile Gln Arg Phe Gly Val Gly Phe Cys Phe Thr Val Val
    450   455   460
Tyr Ala Ala Leu Leu Thr Lys Thr Asn Arg Ile Ala Arg Ile Phe Lys
465   470   475   480
Ala Gly Lys Gln Ser Ala Lys Arg Pro Ser Phe Ile Ser Pro Lys Ser
                485   490   495
Gln Leu Val Ile Cys Ala Cys Leu Val Ser Val Gln Ile Leu Ile Asn
            500   505   510
Gly Val Trp Met Val Ile Ala Pro Ser His Ala Met His His Tyr Pro
        515   520   525
Thr Arg Glu Asp Asn Leu Leu Val Cys Asp Ser Tyr Ile Asp Ala Ser
    530   535   540
Tyr Met Ile Ala Phe Ser Tyr Pro Ile Phe Leu Ile Val Ile Cys Thr
545   550   555   560
Val Tyr Ala Val Leu Thr Arg Lys Ile Pro Glu Ala Phe Asn Glu Ser
                565   570   575
Lys His Ile Gly Phe Thr Met Tyr Thr Thr Cys Val Ile Trp Leu Ala
            580   585   590
Phe Val Pro Leu Tyr Phe Gly Thr Ala Asn His Val Pro Leu Arg Ile
        595   600   605
Thr Ser Met Ser Val Thr Ile Ser Leu Ser Ala Ser Val Thr Ile Ala
    610   615   620
Cys Leu Phe Ser Pro Lys Leu Tyr Ile Ile Leu Ile Arg Pro Glu Arg
625   630   635   640
Asn Val Arg Gln Ser Met Met Pro Pro Arg Tyr Gly Asn Met His Arg
                645   650   655
Thr Ala Gly Thr Gly Pro Ser Ser Met Met Ala Ala Ala Val Val Thr
            660   665   670
Ala Ala Thr Cys Ala Gln Glu Glu Lys Ile Gln Lys His Ile Thr Pro
        675   680   685
Thr Asn Thr Asp Glu Glu Ile Val Arg Asn Gly His Ser Asp Asn Phe
    690   695   700
Gln Tyr Asn His Val Ala Gly His Gln Cys Val Gln Ser Asn Thr Ile
705   710   715   720
Cys Arg Ser Ile Cys Ala Lys Gln Cys His Asn Gly His Asp His Ala
                725   730   735
His Arg

<210> 7
<211> 9400
<212> DNA
<213> Drosophila

<400> 7
cagttagttc ctagttaaca agtgaggttt ggttgaaatc tttctttgct atataggtat 60
ttcatttccg tatgctgtat cggttttgta aaataataaa gcatttcgcg agctctctct 120
ctctctgtat ctcaaaaata taaagtaaat gcccctatgt gtgtgcttaa agcatctaga 180
ttgatgaaaa tgtacacttc aaatgcttcg ctctggtttg cttacaacgt agctgaaatg 240

```
cagcaacaaa atggcaatgt attatttccc tttacttctt atattcctct tgattcgcag 300
ttggaaggaa ttttagttgt ataggtggaa aacggcagct ggatagcaag aaatgacaat 360
ctttcaaaac gtaggatatc tttttggttt atgaaacaca aaacaaatct taagcctgtg 420
tgtatatata tatatatctt acgagcataa tcagcgtaaa tcaaaaacaa ttgatttctt 480
acaaaataca taaaacataa ttcgcctgaa ttctatatat catgtacata cagttcctag 540
atcttataca tatacctaca tccatatata tatacttgat tttctggttg tttaaattag 600
tttttaagta tttaaaaatg gcatgtggtg acatgtatta caatacatca tacatatatg 660
gtacatacat ttatatttaa tatcgatttc caagttggta taaaatgcat gcaagtcgaa 720
gggcccgatt ccaattccaa ttccaacccc ccatcgctcg agttttggtt aagcttgcac 780
ggcaccgcag ggtatggat gtgccctatg gtctaggttt tagttatttc atatacattt 840
ataaagtttt gggattttcg aatcactgca cgtatagcat gacagcatgt tgtttcacac 900
aatgctccat atttcgacta acgactaata ctaggtatag ctagagatcc taagctacaa 960
tttttgccta agactgttcc gggggagttg cgttaactga tcacttagcc agtggttcca 1020
gataacatcg caccgtgttg ttctcatcga gcatggccga caggtggcgc tgctgatggt 1080
gcggaggatg ctgatgggga tgcggatggg ttagctggtg gggatgcgga ttgggattgc 1140
tgtgtgatgt atgcgacgac gatgatgtgg tggactgtgg tgacagtttc tgtagctgct 1200
ggcgcttctg cgcgggcgtc tgattcgccg gcggcagcgg tggtcccgag ttgccgtgac 1260
cgaatagctc atttaccgcc aaacggccgc ctggcgactt ggcattcagt gcttgctgct 1320
gctgctgccg gaatgcgcca cgcagggagc cctcaccaac gccgccgcta accccaaaaa 1380
ccgcaccgtt gtgctcgggt tgcggtcctg ctcctggtgg caaagccccg tagtcaagag 1440
gcgctccatg tttgcgcagc tgacgtctgt aaaggggata aatgtaattt tttcgtaacg 1500
ttgatatagt caaacattag aagtacttac tgaaaccgcc gcttaaaatc ctcgtcctgc 1560
atcgccatca gacgacgcat ccgctcatcg ctctgcggtg gcaatccatg agcatcatcc 1620
tccaagtcgc tctcgtcctc ctcctcatcc tcatcacaga cgctgtcgtg caattgctcc 1680
gcggccgttg cctcctcatc ctcctcctcg tcttcgtcgt catcgtcatc cgccgtgcag 1740
cgtcgagccg atctttgctg agccagcgac tgttgttgct gctgcttctt gagctgctgc 1800
tgttggctcc tggacgagac actgccgcca ctcaggcttc cgccgccgta ccgactaccg 1860
ccggcggcca ccgccgctct catgttgttg aagacgccag tgccaacgct attgttgtac 1920
acctcgtcca catcgaagtc acccacctcg taatgctgca catccgagtt gcggttgctg 1980
aggatctcct cgtagaccgg cccaccatcc gcgccgccga cgctgccttg gttgtagtag 2040
gcatcgccca tgcccagatc acagttggaa tgcttatttc tatactggcg ctgcataaac 2100
tcctgctggt actgcgccat gtgtggcggt ggagcaggga gctgctgctg gctaccagct 2160
cgtcgttgca ggtgcaacag gtgagcatgt tgctcggcct gctcctcgtc cgaggagagt 2220
tgctgctgct gctgctgttg ctgctgctga tgtagctgct ggcgacgcaa ctgttgctgc 2280
tgctggtagt aggcagccgc attgccagtg gctagtggtg ggtccaacgc ctccgcgtat 2340
tgcggatagc ctgttggctg gccaggcttg gtcttcttgt acgatggata gaaaagcacc 2400
tgggaggcgg cggccgtgcg actgtgcaga gatccactag gcgaggcggg aactcctccg 2460
gcgcccctct tctttttgcc attgtggccg ctgttcggct tttgagtcca gtagaatacc 2520
tcgtgcgcct gttgcggcga tgatgcagca gcacccgcat tattgtagta ctcccgcggt 2580
gttttcaatc ctcctgttcc agttgatggt tgcttccagc tgtagtagtc cggctgatgg 2640
tagtaagact cgcccaaagg tccctgttgc tgctgttgtt gttgttgctt gtgcgcaaaa 2700
atcaggtgtg tggacccggc agagctggtg ggtgaggcat gaccagtgga gcggagagtg 2760
ctacctcttt cgagcacaac atggccacca agggtctgcg gtggagcctg tagtgtgccc 2820
gtcattaggc cactttttag ggttcttctt tgctgggcct cgttcagttc acaatcagag 2880
gtggctccca taatggcatc ctctgtgatg ggcactggtg tctgctgttg ctgctgggga 2940
tcgtcatgcg cctcttcgct cactcctaac ggcagcttgg gcttcttgga cgtggagacg 3000
ctgcaaaggg ttccggctgt gggtccattg ccggtaacaa tgcccgagga aattgatgtt 3060
ttcgagcagc gcaagcactt gggcagccaa gcgtgctgtc tcacatactt ccgatactcc 3120
cgccggatct tttcgttttg aatgcaatgg aacacgaaga tatagagccc cagcaatgtg 3180
ttgaagcaga tgaagcagta tccatacaca tcagccgtgt catcgtccat tttggctcct 3240
cgcaaataga agtaggccga gcaccaaaca gcgctgagca gcaagaggaa cacaaaggag 3300
cagcgtatgt cgaacctagg atttgaggta catgggcatc gaaggatctt taaagtgtat 3360
atggattttc aatatcttac cgcacgctag cgaggcgagt atgttccttg gtctttagac 3420
ccgtgcgact tttgcggcac attataatcc aggagaggaa tgtgtaacca atggcagcct 3480
gtaaaagtga attaatttaa aatggtaact gcgcttttga agcttcatct catattatgt 3540
aacttaccac aaagaagaca agcactggta ttacaaaggt ggcataaaac aaggcattcg 3600
cctccatcag cacgcaataa tcgttttggg tataggtgct gggatcgatg accagcgaga 3660
tggccaccac gctcagcgac agtccgtagg acaagaggta gtagcagttc accttgggcg 3720
tctggtccac ctccagcagg agctcgtccg aggtgagcgt tgagtacgaa tggaaggctg 3780
cagaggcaaa gggttagggc tgcaaagtgt agatgaaatt cggccttctt accttcgtaa 3840
```

cagaaccagg cggtgcccga taggatggca cagtgtagga aaatagtaat gaagccgcag 3900
aaaatgcttg tttcggtctg ttcaatgccc aggagaaaga gcagctcgat ggccaggagg 3960
caaaggtaaa tgctggtata gatcgaggtg cgcgcggact ggagtggaca gttaaagaga 4020
agacagagtg aaaagttaat ataagttcga gaagtcagtg caaatttttt ctttaagcag 4080
caccagcaga aggaccctaa agattcgatc atcagggagc ttttagacat tttaaggaaa 4140
ttgataatct taaacgtttt gaagaatagg tgtcagattt gatagtcaaa cttaatttag 4200
aacttgattt acaataattt tggtgggtgg tggtgaaaat ccaaagacta agagtaatac 4260
gaatacgaat aaaacaatga acagaacacg aaacaagcac aacacatagg taatatacat 4320
acaaactaaa tgaacaaaca caaaatgcaa tatttacagg acaaaattag caacatttgt 4380
atattttatt aattaatgaa atacaaaaga aagtcagaag tgttctaggt atttagttgt 4440
acgatgtcta tgttaaattt gttaaacaat caatgcacat aaaaggggtt atggtatttc 4500
aaaaatgatt aaacgattaa gtaaaacatg ccagagtgat gatttgtaac aaaaagaaaa 4560
atgagtctat catatgaaga tatttcatta ttggttttgt gtgctaatct taagctatga 4620
caaattacta cgatacataa atataaatgg gagtcttcgc gtgcaagctg ttactgggac 4680
gcaacagtga tcatcgctgc ccagcacagc ttttgagtct tatctttgaa tcctgccaga 4740
gagaggttca ttaacaattt aagttactta agcattatat cacatgtgtc tctcggtcat 4800
aaatacatat gtctagaaag cctagcctag cctaaaattg tgtttatctg ttttttcacc 4860
accaccacca cttatattgt aagtcgtcca tgttggcctg cataatgatg ttgttcgtgt 4920
tatccaatcc attgccgtcc acattaaggt tgtgattgcg gtgcggatgt ggctgcggct 4980
gctggccacg ctgctgctgc tgctgctgct ggacatgctg ctgctgcaga tagttgtggc 5040
tgatgttgag gttattgctg gcgttgctgg cctgaactgc atgttgctgc agtgcagcgg 5100
ctattgaggc agcagcagct gcagcagcag ccgtctgctg accgtgctga tgcagattca 5160
gattcaagtt gttcatggca tgttgcatat ttcgctgcgg ctgattggac agcacaacag 5220
cagcagcagc tgccgcaaca acttgttgtt gctgttgctg ctgctgctgt tggacattgt 5280
agttcagatt gttgcggtgt gagttgcgga tggaattgtg ttggataaag ttcgtgttcc 5340
catgactggg cggcggcaca ttggactgac tggttggcgt ggtcagggtc aaggactcgt 5400
gggtctggtc gcgaatattg ttttgctggc gtctgctgcc cgaccgctga tgcggcaagg 5460
gattggagcc gtttcttacc ttcacaaaga ccccattgaa cagcttcagc gttagcaggg 5520
cgataactat gaagaccacg cagatgccga tgcttatgta gatgaatatg cgcatgtttc 5580
catcgaacat ggtgaacaac gaatgctggt gctcatccac aacgtccatt agtatggcaa 5640
agtttgtcag gtggttgcaa ctgcagacgc tgtgcgtgcg gttagtggac tccagactgc 5700
atccgttggc agaccacgca ctaaaacaaa aacattaatt cgtataaatt attaggcata 5760
tgagcatatg agtcttactg gtcaatatag ttccagaaca cgcaggtggg attcgttaca 5820
ttctcggtct tcagatgttt cagtgtcagg gttatgggct gggagagttg tatgtgacgc 5880
cccttgccca agctggccga gatcaccttg ctgttcagga tgcgtaggcg ctgttggatc 5940
tcccccgcgt tgacatcgct gtcgttgctc aggatggctg tattgcgaac taagaaaaca 6000
tttagattat actattctaa cgtttttcag ttctctgtgg acttaccgta actgcgggag 6060
ctcttgagat cgaagtgatc atagctgggc tttagaatgg attccaggcg atcgaaggcg 6120
gcgaatacaa ttcgcaccag accgccttca ctattatcta ttagagcagc tcgtggcagc 6180
tcaatacgat ccgaactcaa gggccactga tccgaatctg ggaagaccac gctggactgg 6240
atagtcttgg tttccagaac tcgaacggag agcactatgg tgatatacag ttataagtat 6300
tttgaaatgc aaacaatgca aacatttaga actcacatat atttttgact ttttgcacca 6360
cgctgcgctc cctgatgatc gtatccgcca gcagaaaggc attgtattcc aggccagtta 6420
gcaaggatgt ggctacacgc atttggtcct ccggattgag atccaaccac gaggacagct 6480
gcgattcgtc cagcaggttg gagccggttt tgaccacaca atgcaacaac tccatgatca 6540
tagcctcgcg ctgtcgctga tccgggaagg tctccttgtc gtgcatcatc ttctcggaca 6600
ctgtttggat aatcttcgtg gtgaccaaca tgtcgccgcc gtagagcgtt ttgctactgg 6660
tcacttcgga catatcgttg gcgatggaga tcaaggagga gtcccgctgg tttactcgca 6720
tttccagatt gttgagccaa aggctgcggc attgcgtcag atcaggtgtc gcgggatgcc 6780
aggtgggctc aaagttgcgc agacgctggt ttaccttgtg ggccatgctc accggcggct 6840
cgcagctgct gctgttgtgc agacagtcgc caccgctggg tgccggagtt gtcgaactga 6900
tgtcatcatc gtactcatcg tagccggagt cgggtatcct cttcattaga acgcaacgcc 6960
acttggcgat gccggctgct ccgccaggac aggctgaac attcacatca cccactcgag 7020
tcatgttcca gtacaaattg cgggcatgag tggggccgca aaacaacttg gactcatcat 7080
cgcctccgga catcccggct gcgttcaaat taatattgtt gatggtgcgc accacgctgc 7140
cgttcccggg gacaacggat ccactggccg cagtgcccgg tgggcggttg ggtgagctct 7200
ttgtggtcag tagggttccg tcatcagttc cggaacctcc gacgcccgtg agtatcctcg 7260
tgttggatgg tgaagtggca ttaccgccag ccggcagctt gctgctcggc ttggtgggtg 7320
aggcagtgtc ttcaccgtgg tggttgtggt ggtggtgcag ttgcggcggc ggtggcagac 7380
catcatggcg gccagccgga tgcttggtgg tcgtcgtgga ggtggcccca cccttcaaac 7440

```
gtccaccgcc cacggctgtc gtgctgctcg agggcgtgga gtgcgtgacg gcggtgtgtt 7500
gcggtggtac gttgaacagg ccgggattcc cactgattcc caccactccg ggaagtgtgg 7560
gaagtctcgg cgggggcggc gcacccgctc caaccccggg cggatggatc agtccactgc 7620
cgttgccaaa gatcggcgga ccattgctca gcacccatgg cggcggactg ggcctgttgg 7680
tcgtcgtcga agtttgggct gcacttatgc actggtagtg tgcctccaga tacttgtggg 7740
taccgggaca gggatccccg aacatgctcg tggctgccaa cacgccgcag ctctgcttgt 7800
gggcacatct ggaaaatggg aaaaatacgc gttcatcagt aaacagagtt gggtaatttt 7860
cacatattta agatgcttga ttgattgact tgtgttgaaa tagtttattg atattcgcaa 7920
ccagagtatg aaatttccaa atgacatgtt taatgatttg aagatccaat aaactttaaa 7980
agatgtatta ccttgagttc agtacgctga gtgacttggg aaacatgcag ttaacactcc 8040
actccacatt cccgtggtca ttgcagatgg taatcgagaa gcggccatag ttggccccgaa 8100
tgaggttgat cacatcgccg ggatcgcact cgatggtcag tttcttacct tcgcaggcgt 8160
aggcggtttg gtacttggac actgaaaggc gaaggcttca taagtcacag ggtgctgggc 8220
agggtatggg tttatatgga catggtatat atatatggca tataaaaaga aatttttttgg 8280
ttagcggcga tgattgttgt agtttggttt ggtttggatt ggattgttgg tgggcctgca 8340
aagacactta cgatctaacg atatatgctc ataggatatt gacaatatgg ttggtagcat 8400
gatacgcttt tacactgcct gggatatact tagtaatgct gaggttgatg gtggagtatg 8460
tgggtcgtta gcttggtagc aagtgcaata gtagtcataa ataccaaatc ctgccggccc 8520
agatcacgaa tcatcgatcg tccatcattg gttcttatcg gcaatgctca ttcggcagtg 8580
tttcactcgg tttgaagtat tcaaactgaa agggaagata acatttataa tgtttggtat 8640
tattattatt attattttaa ggcacataac taataactta agacatgttg tatttcgtta 8700
gaaaggaaaa ggagaactat ccttttcttt tattgatatt ggttccattt aatttacttg 8760
tgcaatatca catcaatgtt gagtcaatct aatgacctgg ttacagtaaa tatcaaaagt 8820
gtagattcta ataaatttaa gttattttaa tattcagaac ctctttgtat aactcctatt 8880
aaaacactat ataagcaatg caaatatcaa gaaaacataa ttaattgtaa gaattataaa 8940
ttgcatttat atatctttaa taattaaaat ggtcttatat ttatgtatga tgatatccag 9000
gccaagcaca cctatgtaca tatgtacata cctcttatgt atttccgttt ggagtcatat 9060
tcatgaaagt tggactcaaa agtaccattc aacttatgta tctgaaaggt agagaatcca 9120
ttagataatt gtttacatat tgactccttg ggcaaacacg aacaagcagg caaaagacaa 9180
atagaagata agaagatggc atggggcagg ggaggggggta gggtagggag ctaagaatac 9240
acacccttct gtgtacaccc acactgtggt tgtagttgtt gcatggccat cgattcgtgg 9300
acagaacaaa gacccgtttt ctttgtacac ccgatgcaca ctattcacga acatatatgc 9360
atgtggatgt aaatgtagag aaaagttgct gcagctgcgc                       9400
```

<210> 8
<211> 5718
<212> DNA
<213> Drosophila

<400> 8

```
atgctaccaa ccatattgtc aatatcctat gagcatatat cgttagatct gtccaagtac 60
caaaccgcct acgcctgcga aggtaagaaa ctgaccatcg agtgcgatcc cggcgatgtg 120
atcaacctca ttcgggccaa ctatggccgc ttctcgatta ccatctgcaa tgaccacggg 180
aatgtggagt ggagtgttaa ctgcatgttt cccaagtcac tcagcgtact gaactcaaga 240
tgtgcccaca agcagagctg cggcgtgttg gcagccacga gcatgttcgg ggatccctgt 300
cccggtaccc acaagtatct ggaggcacac taccagtgca taagtgcagc ccaaacttcg 360
acgacgacca acaggcccag tccgccgcca tgggtgctga gcaatggtcc gccgatcttt 420
ggcaacggca gtggactgat ccatccgccc ggggttggag cgggtgcgcc gcccccgccg 480
agacttccca cacttcccgg agtggtggga atcagtggga atcccggcct gttcaacgta 540
ccaccgcaac acaccgccgt cacgcactcc acgccctcga gcagcacgac agccgtgggc 600
ggtggacgtt tgaagggtgg ggccacctcc acgacgacca ccaagcatcc ggctggccgc 660
catgatggtc tgccaccgcc gccgcaactg caccaccacc acaaccacca cggtgaagac 720
actgcctcac ccaccaagcc gagcagcaag ctgccggctg gcggtaatgc cacttcacca 780
tccaacacga ggatactcac gggcgtcgga ggttccggaa ctgatgacgg aaccctactg 840
accacaaaga gctcacccaa ccgcccaccg ggcactgcgg ccagtggatc cgttgtcccc 900
gggaacggca gcgtggtgcg caccatcaac aatattaatt tgaacgcagc cgggatgtcc 960
ggaggcgatg atgagtccaa gttgttttgc ggccccactc atgcccgcaa tttgtactgg 1020
aacatgactc gagtgggtga tgtgaatgtt cagccctgtc ctggcggagc agccggcatc 1080
```

gccaagtggc gttgcgttct aatgaagagg atacccgact ccggctacga tgagtacgat 1140
gatgacatca gttcgacaac tccggcaccc agcggtggcg actgtctgca caacagcagc 1200
agctgcgagc cgccggtgag catggcccac aaggtaaacc agcgtctgcg caactttgag 1260
cccacctggc atcccgcgac acctgatctg acgcaatgcc gcagcctttg gctcaacaat 1320
ctggaaatgc gagtaaacca gcgggactcc tccttgatct ccatcgccaa cgatatgtcc 1380
gaagtgacca gtagcaaaac gctctacggc ggcgacatgt tggtcaccac gaagattatc 1440
caaacagtgt ccgagaagat gatgcacgac aaggagacct tcccggatca gcgacagcgc 1500
gaggctatga tcatggagtt gttgcattgt gtggtcaaaa ccggctccaa cctgctggac 1560
gaatcgcagc tgtcctcgtg gttggatctc aatccggagg accaaatgcg tgtagccaca 1620
tccttgctaa ctggcctgga atacaatgcc tttctgctgg cggatacgat catcagggag 1680
cgcagcgtgg tgcaaaaagt caaaaatata ttgctctccg ttcgagttct ggaaaccaag 1740
actatccagt ccagcgtggt cttcccagat tcggatcagt ggcccttgag ttcggatcgt 1800
attgagctgc cacgagctgc tctaatagat aatagtgaag gcggtctggt gcgaattgta 1860
ttcgccgcct tcgatcgcct ggaatccatt ctaaagccca gctatgatca cttcgatctc 1920
aagagctccc gcagttacgc catcctgagc aacgacagcg atgtcaacgc gggggagatc 1980
caacagcgcc tacgcatcct gaacagcaag gtgatctcgg ccagcttggg caaggggcgt 2040
cacatacaac tctcccagcc cataaccctg acactgaaac atctgaagac cgagaatgta 2100
acgaatccca cctgcgtgtt ctggaactat attgaccatg cgtggtctgc caacggatgc 2160
agtctggagt ccactaaccg cacgcacagc gtctgcagtt gcaaccacct gacaaacttt 2220
gccatactaa tggacgttgt ggatgagcac cagcattcgt tgttcaccat gttcgatgga 2280
aacatgcgca tattcatcta cataagcatc ggcatctgcg tggtcttcat agttatcgcc 2340
ctgctaacgc tgaagctgtt caatggggtc tttgtgaagg taagaaacgg ctccaatccc 2400
ttgccgcatc agcggtcggg cagcagacgc cagcaaaaca atattcgcga ccagacccac 2460
gagtccttga ccctgaccac gccaaccagt cagtccaatg tgccgccgcc cagtcatggg 2520
aacacgaact ttatccaaca caattccatc cgcaactcac accgcaacaa tctgaactac 2580
aatgtccaac agcagcagca gcaacagcaa caacaagttg ttgcggcagc tgctgctgct 2640
gttgtgctgt ccaatcagcc gcagcgaaat atgcaacatg ccatgaacaa cttgaatctg 2700
aatctgcatc agcacggtca gcagacggct gctgctgcag ctgctgctgc ctcaatagcc 2760
gctgcactgc agcaacatgc agttcaggcc agcaacgcca gcaataacct caacatcagc 2820
cacaactatc tgcagcagca gcatgtccag cagcagcagc agcagcagcg tggccagcag 2880
ccgcagccac atccgcaccg caatcacaac cttaatgtgg acggcaatgg attggataac 2940
acgaacaaca tcattatgca ggccaacatg gacgacttac aatataagtg gtggtggtgc 3000
atttaccttt gcctcctggc catcgagctg ctctttctcc tgggcattga acagaccgaa 3060
acaagcattt tctgcggctt cattactatt ttcctacact gtgccatcct atcgggcacc 3120
gcctggttct gttacgaagc cttccattcg tactcaacgc tcacctcgga cgagctcctg 3180
ctggaggtgg accagacgcc caaggtgaac tgctactacc tcttgtccta cggactgtcg 3240
ctgagcgtgg tggccatctc gctggtcatc gatcccagca cctataccca aaacgattat 3300
tgcgtgctga tggaggcgaa tgccttgttt tatgccacct ttgtaatacc agtgcttgtc 3360
ttctttgtgg ctgccattgg ttacacattc ctctcctgga ttataatgtg ccgcaaaagt 3420
cgcacgggtc taaagaccaa ggaacatact cgcctcgcta gcgtgcggtt cgacatacgc 3480
tgctcctttg tgttcctctt gctgctcagc gctgtttggt gctcggccta cttctatttg 3540
cgaggagcca aaatggacga tgacacggct gatgtgtatg gatactgctt catctgcttc 3600
aacacattgc tggggctcta tatcttcgtg ttccattgca ttcaaaacga aaagatccgg 3660
cgggagtatc ggaagtatgt gagacagcac gcttggctgc ccaagtgctt gcgctgctcg 3720
aaaacatcaa tttcctcggg cattgttacc ggcaatggac ccacagccgg aaccctttgc 3780
agcgtctcca cgtccaagaa gcccaagctg ccgttaggag tgagcgaaga ggcgcatgac 3840
gatccccagc agcaacagca gacaccagtg cccatcacag aggatgccat tatgggagcc 3900
acctctgatt gtgaactgaa cgaggcccag caaagaagaa ccctaaaaag tggcctaatg 3960
acgggcacac tacaggctcc accgcagacc cttggtggcc atgttgtgct cgaaagaggt 4020
agcactctcc gctccactgg tcatgcctca cccaccagct ctgccgggtc cacacacctg 4080
atttttgcgc acaagcaaca acaacaacag cagcaacagg gacctttggg cgagtcttac 4140
taccatcagc cggactacta cagctggaag caaccatcaa ctggaacagg aggattgaaa 4200
acaccgcggg agtactacaa taatgcgggt gctgctgcat catcgccgca acaggcgcac 4260
gaggtattct actggactca aaagccgaac agcggccaca atggcaaaaa gaagaggggc 4320
gccggaggag ttcccgcctc gcctagtgga tctctgcaca gtcgcacggc cgccgcctcc 4380
caggtgcttt tctatccatc gtacaagaag accaagcctg gccagccaac aggctatccg 4440
caatacgcgg aggcgttgga cccaccacta gccactggca atgcggctgc ctactaccag 4500
cagcagcaac agttgcgtcg ccagcagcta catcagcagc agcaacagca gcagcagcag 4560
caactctcct cggacgagga gcaggccgag caacatgctc acctgttgca cctgcaacga 4620
cgagctggta gccagcagca gctccctgct ccaccgccac acatggcgca gtaccagcag 4680

gagtttatgc agcgccagta tagaaataag cattccaact gtgatctggg catgggcgat 4740
gcctactaca accaaggcag cgtcggcggc gcggatggtg ggccggtcta cgaggagatc 4800
ctcagcaacc gcaactcgga tgtgcagcat tacgaggtgg gtgacttcga tgtggacgag 4860
gtgtacaaca atagcgttgg cactggcgtc ttcaacaaca tgagagcggc ggtggccgcc 4920
ggcggtagtc ggtacggcgg cggaagcctg agtggcggca gtgtctcgtc caggagccaa 4980
cagcagcagc tcaagaagca gcagcaacaa cagtcgctgg ctcagcaaag atcggctcga 5040
cgctgcacgg cggatgacga tgacgacgaa gacgaggagg aggatgagga ggcaacggcc 5100
gcggagcaat tgcacgacag cgtctgtgat gaggatgagg aggaggacga gagcgacttg 5160
gaggatgatg ctcatggatt gccaccgcag agcgatgagc ggatgcgtcg tctgatggcg 5220
atgcaggacg aggattttaa gcggcggttt caacgtcagc tgcgcaaaca tggagcgcct 5280
cttgactacg ggctttgcc accaggagca ggaccgcaac ccgagcacaa cggtgcggtt 5340
tttggggtta gcggcggcgt tggtgagggc tccctgcgtg gcgcattccg gcagcagcag 5400
cagcaagcac tgaatgccaa gtcgccaggc ggccgtttgg cggtaaatga gctattcggt 5460
cacggcaact cgggaccacc gctgccgccg gcgaatcaga cgcccgcgca gaagcgccag 5520
cagctacaga aactgtcacc acagtccacc acatcatcgt cgtcgcatac atcacacagc 5580
aatcccaatc cgcatcccca ccagctaacc catccgcatc cccatcagca tcctccgcac 5640
catcagcagc gccacctgtc ggccatgctc gatgagaaca acacggtgcg atgttatctg 5700
gaaccactgg ctaagtga 5718

<210> 9
<211> 1905
<212> PRT
<213> Drosophila

<400> 9
Met Leu Pro Thr Ile Leu Ser Ile Ser Tyr Glu His Ile Ser Leu Asp
1 5 10 15
Leu Ser Lys Tyr Gln Thr Ala Tyr Ala Cys Glu Gly Lys Lys Leu Thr
20 25 30
Ile Glu Cys Asp Pro Gly Asp Val Ile Asn Leu Ile Arg Ala Asn Tyr
35 40 45
Gly Arg Phe Ser Ile Thr Ile Cys Asn Asp His Gly Asn Val Glu Trp
50 55 60
Ser Val Asn Cys Met Phe Pro Lys Ser Leu Ser Val Leu Asn Ser Arg
65 70 75 80
Cys Ala His Lys Gln Ser Cys Gly Val Leu Ala Ala Thr Ser Met Phe
85 90 95
Gly Asp Pro Cys Pro Gly Thr His Lys Tyr Leu Glu Ala His Tyr Gln
100 105 110
Cys Ile Ser Ala Ala Gln Thr Ser Thr Thr Thr Asn Arg Pro Ser Pro
115 120 125
Pro Pro Trp Val Leu Ser Asn Gly Pro Pro Ile Phe Gly Asn Gly Ser
130 135 140
Gly Leu Ile His Pro Pro Gly Val Gly Ala Gly Ala Pro Pro Pro Pro
145 150 155 160
Arg Leu Pro Thr Leu Pro Gly Val Val Gly Ile Ser Gly Asn Pro Gly
165 170 175
Leu Phe Asn Val Pro Pro Gln His Thr Ala Val Thr His Ser Thr Pro
180 185 190
Ser Ser Ser Thr Thr Ala Val Gly Gly Gly Arg Leu Lys Gly Gly Ala
195 200 205
Thr Ser Thr Thr Thr Thr Lys His Pro Ala Gly Arg His Asp Gly Leu
210 215 220
Pro Pro Pro Pro Gln Leu His His His His Asn His His Gly Glu Asp
225 230 235 240
Thr Ala Ser Pro Thr Lys Pro Ser Ser Lys Leu Pro Ala Gly Gly Asn
245 250 255
Ala Thr Ser Pro Ser Asn Thr Arg Ile Leu Thr Gly Val Gly Gly Ser

```
            260                 265                 270
Gly Thr Asp Asp Gly Thr Leu Leu Thr Thr Lys Ser Ser Pro Asn Arg
        275                 280                 285
Pro Pro Gly Thr Ala Ala Ser Gly Ser Val Val Pro Gly Asn Gly Ser
    290                 295                 300
Val Val Arg Thr Ile Asn Asn Ile Asn Leu Asn Ala Ala Gly Met Ser
305                 310                 315                 320
Gly Gly Asp Asp Glu Ser Lys Leu Phe Cys Gly Pro Thr His Ala Arg
                325                 330                 335
Asn Leu Tyr Trp Asn Met Thr Arg Val Gly Asp Val Asn Val Gln Pro
            340                 345                 350
Cys Pro Gly Gly Ala Ala Gly Ile Ala Lys Trp Arg Cys Val Leu Met
        355                 360                 365
Lys Arg Ile Pro Asp Ser Gly Tyr Asp Glu Tyr Asp Asp Asp Ile Ser
    370                 375                 380
Ser Thr Thr Pro Ala Pro Ser Gly Gly Asp Cys Leu His Asn Ser Ser
385                 390                 395                 400
Ser Cys Glu Pro Pro Val Ser Met Ala His Lys Val Asn Gln Arg Leu
                405                 410                 415
Arg Asn Phe Glu Pro Thr Trp His Pro Ala Thr Pro Asp Leu Thr Gln
            420                 425                 430
Cys Arg Ser Leu Trp Leu Asn Asn Leu Glu Met Arg Val Asn Gln Arg
        435                 440                 445
Asp Ser Ser Leu Ile Ser Ile Ala Asn Asp Met Ser Glu Val Thr Ser
    450                 455                 460
Ser Lys Thr Leu Tyr Gly Gly Asp Met Leu Val Thr Thr Lys Ile Ile
465                 470                 475                 480
Gln Thr Val Ser Glu Lys Met Met His Asp Lys Glu Thr Phe Pro Asp
                485                 490                 495
Gln Arg Gln Arg Glu Ala Met Ile Met Glu Leu Leu His Cys Val Val
            500                 505                 510
Lys Thr Gly Ser Asn Leu Leu Asp Glu Ser Gln Leu Ser Ser Trp Leu
        515                 520                 525
Asp Leu Asn Pro Glu Asp Gln Met Arg Val Ala Thr Ser Leu Leu Thr
    530                 535                 540
Gly Leu Glu Tyr Asn Ala Phe Leu Leu Ala Asp Thr Ile Ile Arg Glu
545                 550                 555                 560
Arg Ser Val Val Gln Lys Val Lys Asn Ile Leu Leu Ser Val Arg Val
                565                 570                 575
Leu Glu Thr Lys Thr Ile Gln Ser Ser Val Val Phe Pro Asp Ser Asp
            580                 585                 590
Gln Trp Pro Leu Ser Ser Asp Arg Ile Glu Leu Pro Arg Ala Ala Leu
        595                 600                 605
Ile Asp Asn Ser Glu Gly Gly Leu Val Arg Ile Val Phe Ala Ala Phe
    610                 615                 620
Asp Arg Leu Glu Ser Ile Leu Lys Pro Ser Tyr Asp His Phe Asp Leu
625                 630                 635                 640
Lys Ser Ser Arg Ser Tyr Ala Ile Leu Ser Asn Asp Ser Asp Val Asn
                645                 650                 655
Ala Gly Glu Ile Gln Gln Arg Leu Arg Ile Leu Asn Ser Lys Val Ile
            660                 665                 670
Ser Ala Ser Leu Gly Lys Gly Arg His Ile Gln Leu Ser Gln Pro Ile
        675                 680                 685
Thr Leu Thr Leu Lys His Leu Lys Thr Glu Asn Val Thr Asn Pro Thr
    690                 695                 700
Cys Val Phe Trp Asn Tyr Ile Asp His Ala Trp Ser Ala Asn Gly Cys
705                 710                 715                 720
Ser Leu Glu Ser Thr Asn Arg Thr His Ser Val Cys Ser Cys Asn His
                725                 730                 735
Leu Thr Asn Phe Ala Ile Leu Met Asp Val Val Asp Glu His Gln His
```

```
            740                 745                 750
Ser Leu Phe Thr Met Phe Asp Gly Asn Met Arg Ile Phe Ile Tyr Ile
        755                 760                 765
Ser Ile Gly Ile Cys Val Val Phe Ile Val Ile Ala Leu Leu Thr Leu
    770                 775                 780
Lys Leu Phe Asn Gly Val Phe Val Lys Val Arg Asn Gly Ser Asn Pro
785                 790                 795                 800
Leu Pro His Gln Arg Ser Gly Ser Arg Arg Gln Gln Asn Asn Ile Arg
                805                 810                 815
Asp Gln Thr His Glu Ser Leu Thr Leu Thr Thr Pro Thr Ser Gln Ser
            820                 825                 830
Asn Val Pro Pro Pro Ser His Gly Asn Thr Asn Phe Ile Gln His Asn
        835                 840                 845
Ser Ile Arg Asn Ser His Arg Asn Asn Leu Asn Tyr Asn Val Gln Gln
    850                 855                 860
Gln Gln Gln Gln Gln Gln Gln Gln Val Val Ala Ala Ala Ala Ala Ala
865                 870                 875                 880
Val Val Leu Ser Asn Gln Pro Gln Arg Asn Met Gln His Ala Met Asn
                885                 890                 895
Asn Leu Asn Leu Asn Leu His Gln His Gly Gln Gln Thr Ala Ala Ala
            900                 905                 910
Ala Ala Ala Ala Ala Ser Ile Ala Ala Ala Leu Gln Gln His Ala Val
        915                 920                 925
Gln Ala Ser Asn Ala Ser Asn Asn Leu Asn Ile Ser His Asn Tyr Leu
    930                 935                 940
Gln Gln Gln His Val Gln Gln Gln Gln Gln Gln Gln Arg Gly Gln Gln
945                 950                 955                 960
Pro Gln Pro His Pro His Arg Asn His Asn Leu Asn Val Asp Gly Asn
                965                 970                 975
Gly Leu Asp Asn Thr Asn Asn Ile Ile Met Gln Ala Asn Met Asp Asp
            980                 985                 990
Leu Gln Tyr Lys Trp Trp Trp Cys Ile Tyr Leu Cys Leu Leu Ala Ile
        995                 1000                1005
Glu Leu Leu Phe Leu Leu Gly Ile Glu Gln Thr Glu Thr Ser Ile Phe
    1010                1015                1020
Cys Gly Phe Ile Thr Ile Phe Leu His Cys Ala Ile Leu Ser Gly Thr
1025                1030                1035                1040
Ala Trp Phe Cys Tyr Glu Ala Phe His Ser Tyr Ser Thr Leu Thr Ser
                1045                1050                1055
Asp Glu Leu Leu Leu Glu Val Asp Gln Thr Pro Lys Val Asn Cys Tyr
            1060                1065                1070
Tyr Leu Leu Ser Tyr Gly Leu Ser Leu Ser Val Val Ala Ile Ser Leu
        1075                1080                1085
Val Ile Asp Pro Ser Thr Tyr Thr Gln Asn Asp Tyr Cys Val Leu Met
    1090                1095                1100
Glu Ala Asn Ala Leu Phe Tyr Ala Thr Phe Val Ile Pro Val Leu Val
1105                1110                1115                1120
Phe Phe Val Ala Ala Ile Gly Tyr Thr Phe Leu Ser Trp Ile Ile Met
                1125                1130                1135
Cys Arg Lys Ser Arg Thr Gly Leu Lys Thr Lys Glu His Thr Arg Leu
            1140                1145                1150
Ala Ser Val Arg Phe Asp Ile Arg Cys Ser Phe Val Phe Leu Leu Leu
        1155                1160                1165
Leu Ser Ala Val Trp Cys Ser Ala Tyr Phe Tyr Leu Arg Gly Ala Lys
    1170                1175                1180
Met Asp Asp Asp Thr Ala Asp Val Tyr Gly Tyr Cys Phe Ile Cys Phe
1185                1190                1195                1200
Asn Thr Leu Leu Gly Leu Tyr Ile Phe Val Phe His Cys Ile Gln Asn
                1205                1210                1215
Glu Lys Ile Arg Arg Glu Tyr Arg Lys Tyr Val Arg Gln His Ala Trp
```

```
            1220                    1225                    1230
Leu Pro Lys Cys Leu Arg Cys Ser Lys Thr Ser Ile Ser Ser Gly Ile
        1235                    1240                    1245
Val Thr Gly Asn Gly Pro Thr Ala Gly Thr Leu Cys Ser Val Ser Thr
    1250                    1255                    1260
Ser Lys Lys Pro Lys Leu Pro Leu Gly Val Ser Glu Glu Ala His Asp
1265                    1270                    1275                    1280
Asp Pro Gln Gln Gln Gln Gln Thr Pro Val Pro Ile Thr Glu Asp Ala
                1285                    1290                    1295
Ile Met Gly Ala Thr Ser Asp Cys Glu Leu Asn Glu Ala Gln Gln Arg
            1300                    1305                    1310
Arg Thr Leu Lys Ser Gly Leu Met Thr Gly Thr Leu Gln Ala Pro Pro
        1315                    1320                    1325
Gln Thr Leu Gly Gly His Val Val Leu Glu Arg Gly Ser Thr Leu Arg
    1330                    1335                    1340
Ser Thr Gly His Ala Ser Pro Thr Ser Ser Ala Gly Ser Thr His Leu
1345                    1350                    1355                    1360
Ile Phe Ala His Lys Gln Gln Gln Gln Gln Gln Gln Gln Gly Pro Leu
                1365                    1370                    1375
Gly Glu Ser Tyr Tyr His Gln Pro Asp Tyr Tyr Ser Trp Lys Gln Pro
            1380                    1385                    1390
Ser Thr Gly Thr Gly Gly Leu Lys Thr Pro Arg Glu Tyr Tyr Asn Asn
        1395                    1400                    1405
Ala Gly Ala Ala Ala Ser Ser Pro Gln Gln Ala His Glu Val Phe Tyr
    1410                    1415                    1420
Trp Thr Gln Lys Pro Asn Ser Gly His Asn Gly Lys Lys Lys Arg Gly
1425                    1430                    1435                    1440
Ala Gly Gly Val Pro Ala Ser Pro Ser Gly Ser Leu His Ser Arg Thr
                1445                    1450                    1455
Ala Ala Ala Ser Gln Val Leu Phe Tyr Pro Ser Tyr Lys Lys Thr Lys
            1460                    1465                    1470
Pro Gly Gln Pro Thr Gly Tyr Pro Gln Tyr Ala Glu Ala Leu Asp Pro
        1475                    1480                    1485
Pro Leu Ala Thr Gly Asn Ala Ala Ala Tyr Tyr Gln Gln Gln Gln Gln
    1490                    1495                    1500
Leu Arg Arg Gln Gln Leu His Gln Gln Gln Gln Gln Gln Gln Gln Gln
1505                    1510                    1515                    1520
Gln Leu Ser Ser Asp Glu Glu Gln Ala Glu Gln His Ala His Leu Leu
                1525                    1530                    1535
His Leu Gln Arg Arg Ala Gly Ser Gln Gln Gln Leu Pro Ala Pro Pro
            1540                    1545                    1550
Pro His Met Ala Gln Tyr Gln Gln Glu Phe Met Gln Arg Gln Tyr Arg
        1555                    1560                    1565
Asn Lys His Ser Asn Cys Asp Leu Gly Met Gly Asp Ala Tyr Tyr Asn
    1570                    1575                    1580
Gln Gly Ser Val Gly Gly Ala Asp Gly Gly Pro Val Tyr Glu Glu Ile
1585                    1590                    1595                    1600
Leu Ser Asn Arg Asn Ser Asp Val Gln His Tyr Glu Val Gly Asp Phe
                1605                    1610                    1615
Asp Val Asp Glu Val Tyr Asn Asn Ser Val Gly Thr Gly Val Phe Asn
            1620                    1625                    1630
Asn Met Arg Ala Ala Val Ala Ala Gly Gly Ser Arg Tyr Gly Gly Gly
        1635                    1640                    1645
Ser Leu Ser Gly Gly Ser Val Ser Ser Arg Ser Gln Gln Gln Gln Leu
    1650                    1655                    1660
Lys Lys Gln Gln Gln Gln Gln Ser Leu Ala Gln Gln Arg Ser Ala Arg
1665                    1670                    1675                    1680
Arg Cys Thr Ala Asp Asp Asp Asp Asp Glu Asp Glu Glu Glu Asp Glu
                1685                    1690                    1695
Glu Ala Thr Ala Ala Glu Gln Leu His Asp Ser Val Cys Asp Glu Asp
```

```
            1700                1705                1710
Glu Glu Glu Asp Glu Ser Asp Leu Glu Asp Asp Ala His Gly Leu Pro
        1715                1720                1725
Pro Gln Ser Asp Glu Arg Met Arg Arg Leu Met Ala Met Gln Asp Glu
    1730                1735                1740
Asp Phe Lys Arg Arg Phe Gln Arg Gln Leu Arg Lys His Gly Ala Pro
1745                1750                1755                1760
Leu Asp Tyr Gly Ala Leu Pro Pro Gly Ala Gly Pro Gln Pro Glu His
                1765                1770                1775
Asn Gly Ala Val Phe Gly Val Ser Gly Gly Val Gly Glu Gly Ser Leu
            1780                1785                1790
Arg Gly Ala Phe Arg Gln Gln Gln Gln Gln Ala Leu Asn Ala Lys Ser
        1795                1800                1805
Pro Gly Gly Arg Leu Ala Val Asn Glu Leu Phe Gly His Gly Asn Ser
    1810                1815                1820
Gly Pro Pro Leu Pro Pro Ala Asn Gln Thr Pro Ala Gln Lys Arg Gln
1825                1830                1835                1840
Gln Leu Gln Lys Leu Ser Pro Gln Ser Thr Thr Ser Ser Ser Ser His
                1845                1850                1855
Thr Ser His Ser Asn Pro Asn Pro His Pro His Gln Leu Thr His Pro
            1860                1865                1870
His Pro His Gln His Pro Pro His His Gln Gln Arg His Leu Ser Ala
        1875                1880                1885
Met Leu Asp Glu Asn Asn Thr Val Arg Cys Tyr Leu Glu Pro Leu Ala
    1890                1895                1900
Lys
1905


<210> 10
<211> 4914
<212> DNA
<213> Drosophila

<400> 10
cgcgggctcc atacgattcg tattagcccg caattgtata aaggttagtt agtttcttag 60
ctagttagtt ggacatttag aagtaggact ttttctgtgc gaaaatcagc agaagttgaa 120
ggaaaaccct cagtaattcc gccaaagttc gttgctggaa aaggactctt ttaaggtaat 180
ttcatgatat tgactcgttt cagttaagtc cctaaagcta atggcacact gaatacaatt 240
aggcattacg gctacgattg caattgcgtt tttttttttc gattgttaag tacacttgtt 300
tcggcatcgg gagaagacag caaagttgga aacctagttt ctggttccca tttttatttt 360
tttgtttttg aagacatgaa catgcgtatc acagtatttg gttgctgcgc gtttctctct 420
ctctctaact caaaaaaaaa gaatttatat aaaaatgcta aatatatata tgtatatatg 480
tgttatgtac gcaaaggcga cacacattaa gtcctaatgt ctaaaaggca agagcgatta 540
ttgaatttgg ttacaccaaa atggtaaaaa ggcaacgcaa ggggcaacta attaggagca 600
acattaagta cacacaggaa aataggatat gaagcccatc tagggcaatt acaaatacat 660
tgtagatgac ctatgacctc gcgataaagg cccaaactaa agcactaaga gcacatttcc 720
ggggttatat acacagtaca tagatcggtt taccccattg tgttctatgg acaggcaaag 780
caactaaaat cgaaataacg atatcgttat ctcacgatcg ttttatgaca aacaaattgg 840
accgtttccc cacccattta agttgctttg cctctttcga tacccgaaac cgcattaagt 900
tatagtagag gattgacttt ggcacagcag ctgctgcagg aacattgact ggaagcagga 960
agcatttcct gccctgaaat ggtacctaat ccaattccgt ctagagtagc gtctcattgc 1020
catactgctg caagtgggga tagtagctgc cgtgctggct gtacttttgc gaggcatagc 1080
tacctggcgg ataatacagc cgatgcgagg atgtggagtg atgctggtga ttgtggtggg 1140
gatggaactg atgatgctgc tgctgctggt gatgatgatg atgcggcgaa tggcgctgct 1200
ggtgatgcac gctcctgccc agcgaatagt tggacatctg attcagcgat atagagccgg 1260
taccaggcga agtcttcagt ttgtacggac ttggcggcaa cggaacactg gccatagaag 1320
ttcgcaccga atggtaggag accaccgtgc tattgcaagc cgcatgcgat ccacgcgagc 1380
tactatgccc gtccaggtgt ccagttgctg ctccggcaga cgtatgcttc accattttgc 1440
cattgctgac ttgcctcttg tagagagtgg aaatgcttgg tccaagtacc caggcaaaac 1500
```

agcaaatgcc cagcatcagt tccatgaagg cgcgcaacaa aaatggccaa agtggcgtat 1560
tggtatcacc cgatcccagc cagacatcga tgttggcaaa ctcgtaaatg acagccaaca 1620
gaagcatggc acatgcaagt ccatagatga aaaggaatat gccaatgccg gaactactgt 1680
gtgcgcttag ctgttgcagt tgctgttgca cgctgagtgc attgctattc ctcaggatct 1740
ccttggtgcg acagtgtacc aaatagccgg agatcagatt catcgagcca aagatccagt 1800
agcaaaaaac gggcgtagcc accaatatct gtagcgcctt atccgactga ttgcctacga 1860
aacaggcgcc tagaagagcg tatggttggg ttactctccg aattctccgg cggatctata 1920

gcatatctca ccaagcaact cgtcagcatc cacgaaacgg gcaacgatca ctgctgacgt 1980
ttgaaaggcc ggcaatcccc aggccacgaa gcacacaaag ttgttttgcg ttagatcctg 2040
ctgagccgtc ttggcactgt tcctcttcgc cggactgcct ccaaagttcg agggaatgac 2100
cacatggccc ttggagtccg gactgtgccg acgaatgtcg cggtgccagc ccaaacagag 2160
cacagcccac ctacaaatag gatatgggaa tatagtttct ataatattaa gtatatatta 2220
taatctaata ccaagcgcag gcggccattc caaaatagta gcgcataagg aagacactag 2280
cgcaggaggc attggacagt ccgtccacgg tgagtaggga ttcattgggc gcctggggat 2340
ccgtgccgca cgcagttccc gttcttccca ccataaagcg cactgcccag cccaaggtga 2400
ccatgttgtg gcaccaaatc agcggcgaca acaatctgga ccacttggcg ctggccaaac 2460
gactgccatc gctggccaaa agacacacgg tggctaccag ggccaaaccc aaagcggcat 2520
aggcccatgt ggagacccag atctctgcca gatgcttctc cgccggtgtg aacagtatgt 2580
ccgcctcgca gagtggtgca cagcgtccag atctgggtag ccttacgtac agatgcgatt 2640
tggccagtcc cgagcaatcc atgggcagtt taccgcccaa tccaccgggt attccttggc 2700
ccggtagccc atacaaatcc tgctcctgct gcggctgatg caattcaccg ggaccttcca 2760
tacacatcgt ctcgtggttg ttctctcgcg gaaatttatc gcagtcgaga gccggtggcc 2820
atggaaatcc gaatccctgc aaaactggat gacaacgtat tcgcacggac tcacagagac 2880
tccggcacgg accgatggca tggactggag ctttgggcgt gcacatgggc acgtaggcgg 2940
cgcacaggaa cagcttcaac tgggaactgc agtcgtattc gatcagtggt gcaaaggttt 3000
gcagtgtgta ctccacgtcc gtttgcattt cgttacccac cagatttggc atggaggtct 3060
cattgtagcc gatcttgcgg cacatctcaa tgcgaatggt ttcgcattgc cggaaggcgg 3120
ggatttctgg cggagaactg gaactggaac tggcactggg gttgccgctc gttgaggatt 3180
ttgagatgcg cggatgcagg aggatcacca ccaggagaca caatatgcag gttggcttca 3240
ttgcggctga tttgttgttc tttcgttccg aagaatacga attcgatacg tttcttcgct 3300
tggtgttccg atttccgttt ctctaaagga atcgaaatgt ggaaaagatc agatttcggg 3360
gttttcttcg cgctgggttc cctcttttct ttccaaagaa tccgaactat atgcagattc 3420
ttcgttcttc ggtggaacga acgctggaag ggtaacctg ttccaggtgc gttcgccctc 3480
tctcgctcgc accctctcac tcgttccaac aacttccagc tctccttctt tccagcttgt 3540
tattgttgct tttggcttgt caactcgtta actacctctt ttttttcggc gtcgttcttg 3600
ttatcgcttc tcctccttcc cccactcgtt cttcagcccc aaccactccc ctttctccac 3660
cgcacgccgc ccaacccgct gatcgacgtc gataaacggg tttttcctca gggattaacc 3720
cgaaaacgat ctagcaaacc atcgaatctt tcacacatct cagtgaacga atagtgtatt 3780
tgtttcgatg tgcttaaacg aatggcgcag tcagtttgta gccttagcga atgtattcta 3840

atctaattgg tattaacatc gaacgtttat atattttatt ttgttttcgt ggtaaaattc 3900
aacgcatact ggtcactctg agctagagat ggctctgcgt cacgtcacct cgcatcgatt 3960
gcaattgctc aacggatctg ctatcgatat caacaagtgt gttggttact aaatgctttt 4020
agatcaaaag gcgccccctg gcggagttat gcgtaatcaa agaacattga gaagggcgcg 4080
cttttcaaaa acaaatttgt ccaaataata agaaacattt aagcataata agaacataaa 4140
tgttcaataa attctatagg tgcagtaaaa tacgtgtttt gctagcacag ctatttaag 4200
ctgagaactt acaatatttt agtacattgt gcttaaggtt taaagataca tatattgaaa 4260
aagttctgga aatgcttttg aagaacaaag ctttaagagg cgcgctttcc gagtatgtga 4320
gctttttca aatggcacac tttctgcgga atataaaata aaaaaaccta ataatatgcc 4380
gcaagaccac ttaagataaa atattaaaaa tattaaaatt gattatattt tgtgcaccaa 4440
aagtggcggt agcttaatat cgcaaaacca aattaatagc aattctgagc caaaccgcga 4500
ttgccatttc tgcatgttgg tcctccgtcc gtttcgccat ctccctttgg ccaccgcttg 4560
gcctctcttt cgcacacaca tggcaaacac aagccacgcc catgtatttc aaccccccttg 4620
acaaaaatac gtaacatgcc acttttatgt atgaaacgtg acaaatacgc acatgtgtga 4680
gaaagagatc caccgagctg gcgagaaaga gagagagaga cgctggacgt ttatgtgttt 4740
ggcaggactc gcgtatgcag tgtgtgtgtg cgtgtgtgtg tgagtgtgtg tgagtgtttg 4800
ttgttggtta acgcacacgc gcagacgtag caacacatac acatatgtgg aaaggatatg 4860
tacatacata catacataca tacatatgcc cgctgacttt ttcgccactg ccaa 4914

<210> 11
<211> 2523
<212> DNA
<213> Drosophila

<400> 11
gaccagtatg cgttgaattt taccacgaaa acaaaataaa atatataaac gttcgatgtt 60
aataccaatt agattagaat acattcgcta aggctacaaa ctgactgcgc cattcgttta 120
agcacatcga aacaaataca ctattcgttc actgagatgt gtgaaagatt cgatggtttg 180
ctagatcgtt ttcgggttaa tccctgagga aaaacccgtt tatcgacgtc gatcagcggg 240
ttgggcggcg tgcggtggag aaaggggagt ggttggggct gaagaacgag tgggggaagg 300
aggagaagcg ataacaagaa cgacgccgaa aaaaaagaga gaaacggaaa tcggaacacc 360
aagcgaagaa acgtatcgaa ttcgtattct tcggaacgaa agaacaacaa atcagccgca 420
atgaagccaa cctgcatatt gtgtctcctg gtggtgatcc tcctgcatcc gcgcatctca 480
aaatcctcaa cgagcggcaa ccccagtgcc agttccagtt ccagttctcc gccagaaatc 540
cccgccttcc ggcaatgcga aaccattcgc attgagatgt gccgcaagat cggctacaat 600
gagacctcca tgccaaatct ggtgggtaac gaaatgcaaa cggacgtgga gtacacactg 660
caaacctttg caccactgat cgaatacgac tgcagttccc agttgaagct gttcctgtgc 720
gccgcctacg tgcccatgtg cacgcccaaa gctccagtcc atgccatcgg tccgtgccgg 780
agtctctgtg agtccgtgcg aatacgttgt catccagttt tgcagggatt cggatttcca 840
tggccaccgg ctctcgactg cgataaattt ccgcgagaga acaaccacga gacgatgtgt 900
atggaaggtc ccggtgaatt gcatcagccg cagcaggagc aggatttgta tggctaccg 960
ggccaaggaa tacccggtgg attgggcggt aaactgccca tggattgctc gggactggcc 1020
aaatcgcatc tgtacgtaag gctacccaga tctggacgct gtgcaccact ctgcgaggcg 1080
gacatactgt tcacaccggc ggagaagcat ctggcagaga tctgggtctc cacatgggcc 1140
tatgccgctt tgggtttggc cctggtagcc accgtgtgtc ttttggccag cgatggcagt 1200
cgtttggcca gcgccaagtg gtccagattg ttgtcgccgc tgatttggtg ccacaacatg 1260
gtcaccttgg gctgggcagt gcgctttatg gtgggaagaa cgggaactgc gtgcggcacg 1320
gatccccagg cgcccaatga atccctactc accgtggacg gactgtccaa tgcctcctgc 1380
gctagtgtct tccttatgcg ctactatttt ggaatggccg cctgcgcttg gtgggctgtg 1440

ctctgtttgg gctggcaccg cgacattcgt cggcacagtc cggactccaa gggccatgtg 1500
gtcattccct cgaactttgg aggcagtccg gcgaagagga acagtgccaa gacggctcag 1560
caggatctaa cgcaaaacaa ctttgtgtgc ttcgtggcct ggggattgcc ggcctttcaa 1620
acgtcagcag tgatcgttgc ccgtttcgtg gatgctgacg agttgcttgg caatcagtcg 1680
gataaggcgc tacagatatt ggtggctacg cccgtttttt gctactggat ctttggctcg 1740
atgaatctga tctccggcta tttggtacac tgtcgcacca aggagatcct gaggaatagc 1800
aatgcactca gcgtgcaaca gcaactgcaa cagctaagcg cacacagtag ttccggcatt 1860
ggcatattcc ttttcatcta tggacttgca tgtgccatgc ttctgttggc tgtcatttac 1920
gagtttgcca acatcgatgt ctggctggga tcgggtgata ccaatacgcc actttggcca 1980
tttttgttgc gcgccttcat ggaactgatg ctgggcattt gctgttttgc ctgggtactt 2040
ggaccaagca tttccactct ctacaagagg caagtcagca atggcaaaat ggtgaagcat 2100
acgtctgccg gagcagcaac tggacacctg gacgggcata gtagctcgcg tggatcgcat 2160
gcggcttgca atagcacggt ggtctcctac cattcggtgc gaacttctat ggccagtgtt 2220
ccgttgccgc caagtccgta caaactgaag acttcgcctg gtaccggctc tatatcgctg 2280
aatcagatgt ccaactattc gctgggcagg agcgtgcatc accagcagcg ccattcgccg 2340
catcatcatc atcaccagca gcagcagcat catcagttcc atccccacca caatcaccag 2400
catcactcca catcctcgca tcggctgtat tatccgccag gtagctatgc ctcgcaaaag 2460
tacagccagc acggcagcta ctatccccac ttgcagcagt atggcaatga gacgctactc 2520
tag 2523

<210> 12
<211> 700
<212> PRT
<213> Drosophila

<400> 12
Met Lys Pro Thr Cys Ile Leu Cys Leu Leu Val Val Ile Leu Leu His
1 5 10 15
Pro Arg Ile Ser Lys Ser Ser Thr Ser Gly Asn Pro Ser Ala Ser Ser

20 25 30
Ser Ser Ser Ser Pro Pro Glu Ile Pro Ala Phe Arg Gln Cys Glu Thr
35 40 45
Ile Arg Ile Glu Met Cys Arg Lys Ile Gly Tyr Asn Glu Thr Ser Met
50 55 60
Pro Asn Leu Val Gly Asn Glu Met Gln Thr Asp Val Glu Tyr Thr Leu
65 70 75 80
Gln Thr Phe Ala Pro Leu Ile Glu Tyr Asp Cys Ser Ser Gln Leu Lys
85 90 95
Leu Phe Leu Cys Ala Ala Tyr Val Pro Met Cys Thr Pro Lys Ala Pro
100 105 110
Val His Ala Ile Gly Pro Cys Arg Ser Leu Cys Glu Ser Val Arg Ile
115 120 125
Arg Cys His Pro Val Leu Gln Gly Phe Gly Phe Pro Trp Pro Pro Ala
130 135 140
Leu Asp Cys Asp Lys Phe Pro Arg Glu Asn Asn His Glu Thr Met Cys
145 150 155 160
Met Glu Gly Pro Gly Glu Leu His Gln Pro Gln Gln Glu Gln Asp Leu
165 170 175
Tyr Gly Leu Pro Gly Gln Gly Ile Pro Gly Gly Leu Gly Gly Lys Leu
180 185 190
Pro Met Asp Cys Ser Gly Leu Ala Lys Ser His Leu Tyr Val Arg Leu
195 200 205
Pro Arg Ser Gly Arg Cys Ala Pro Leu Cys Glu Ala Asp Ile Leu Phe
210 215 220
Thr Pro Ala Glu Lys His Leu Ala Glu Ile Trp Val Ser Thr Trp Ala
225 230 235 240
Tyr Ala Ala Leu Gly Leu Ala Leu Val Ala Thr Val Cys Leu Leu Ala
245 250 255
Ser Asp Gly Ser Arg Leu Ala Ser Ala Lys Trp Ser Arg Leu Leu Ser
260 265 270
Pro Leu Ile Trp Cys His Asn Met Val Thr Leu Gly Trp Ala Val Arg
275 280 285
Phe Met Val Gly Arg Thr Gly Thr Ala Cys Gly Thr Asp Pro Gln Ala
290 295 300
Pro Asn Glu Ser Leu Leu Thr Val Asp Gly Leu Ser Asn Ala Ser Cys
305 310 315 320
Ala Ser Val Phe Leu Met Arg Tyr Tyr Phe Gly Met Ala Ala Cys Ala
325 330 335
Trp Trp Ala Val Leu Cys Leu Gly Trp His Arg Asp Ile Arg Arg His
340 345 350
Ser Pro Asp Ser Lys Gly His Val Val Ile Pro Ser Asn Phe Gly Gly
355 360 365
Ser Pro Ala Lys Arg Asn Ser Ala Lys Thr Ala Gln Gln Asp Leu Thr
370 375 380
Gln Asn Asn Phe Val Cys Phe Val Ala Trp Gly Leu Pro Ala Phe Gln
385 390 395 400
Thr Ser Ala Val Ile Val Ala Arg Phe Val Asp Ala Asp Glu Leu Leu
405 410 415
Gly Asn Gln Ser Asp Lys Ala Leu Gln Ile Leu Val Ala Thr Pro Val
420 425 430
Phe Cys Tyr Trp Ile Phe Gly Ser Met Asn Leu Ile Ser Gly Tyr Leu
435 440 445
Val His Cys Arg Thr Lys Glu Ile Leu Arg Asn Ser Asn Ala Leu Ser
450 455 460
Val Gln Gln Gln Leu Gln Gln Leu Ser Ala His Ser Ser Ser Gly Ile
465 470 475 480
Gly Ile Phe Leu Phe Ile Tyr Gly Leu Ala Cys Ala Met Leu Leu Leu
485 490 495
Ala Val Ile Tyr Glu Phe Ala Asn Ile Asp Val Trp Leu Gly Ser Gly

500 505 510
Asp Thr Asn Thr Pro Leu Trp Pro Phe Leu Leu Arg Ala Phe Met Glu
515 520 525
Leu Met Leu Gly Ile Cys Cys Phe Ala Trp Val Leu Gly Pro Ser Ile
530 535 540
Ser Thr Leu Tyr Lys Arg Gln Val Ser Asn Gly Lys Met Val Lys His
545 550 555 560
Thr Ser Ala Gly Ala Ala Thr Gly His Leu Asp Gly His Ser Ser Ser
565 570 575
Arg Gly Ser His Ala Ala Cys Asn Ser Thr Val Val Ser Tyr His Ser
580 585 590
Val Arg Thr Ser Met Ala Ser Val Pro Leu Pro Pro Ser Pro Tyr Lys
595 600 605
Leu Lys Thr Ser Pro Gly Thr Gly Ser Ile Ser Leu Asn Gln Met Ser
610 615 620
Asn Tyr Ser Leu Gly Arg Ser Val His His Gln Gln Arg His Ser Pro
625 630 635 640
His His His His His Gln Gln Gln Gln His His Gln Phe His Pro His
645 650 655
His Asn His Gln His His Ser Thr Ser Ser His Arg Leu Tyr Tyr Pro
660 665 670
Pro Gly Ser Tyr Ala Ser Gln Lys Tyr Ser Gln His Gly Ser Tyr Tyr
675 680 685
Pro His Leu Gln Gln Tyr Gly Asn Glu Thr Leu Leu
690 695 700

<210> 13
<211> 9795
<212> DNA
<213> Drosophila

<400> 13
tcatactcac tggctgccaa cccctactct actctgacta ctttggttct ggttttggat 60
cttggctgat gctaggttct aactttaagt acgattctta gtattgctca ctggatacac 120
tgtagataca ccgcagatac agccgtcctg cggcggcaca acaacaagaa aacaaacgct 180
aatgcagcta cttacgacta ggaggatgat tagtggcaag gggttaagga gcgggttaat 240
cacaactggc tgggaaactg aaagtggccg acaagcgcat ctacttaacc tagatgatgg 300
tcctacgcgg cccagggtca atcacattgc ggagtaggtt ctacacagag agatgcgatt 360
tggtcaaaaa ctaagcttaa acgaccgcga tgttatctta tcaattaaaa cgaaaaaaat 420
ggaatcaaca tttttataga cccatttaaa gagctatcat ctttttttcac cgaaatactt 480
aaggaaataa tatgcatgct tagtaaagga acactacact aggtcttaag aaatgattat 540
caagaagtaa acatttggca ccccgaggcg aatactgata ctggtagttt ttagaaaact 600
ttgctgatac cgaaaactcg cgaattaaga taaattctaa cattgcacaa atttcttagt 660
ttacattgta aattttgtct acttctaggc agagcacaat gtgcctgcaa ttttctccgt 720
gttgttgatt ctgctaagtt caggaaaatg actagctatg tgaacttgac ttgaggaggg 780
agggggaggt tcgtatcgtg acttatctag cgttttaccc taatgcattt gattgcttta 840
agctgatgat tggtcagttg cacacgcgct attctggggc agattttgtg ggcgtgtcgt 900
aaaataattg cttggtgttt cgcttctaat tgacgaggac ttggacttgg atttgaagtg 960
tggaccgtgt acagagtgga caatggacag ggcatcttat ttattcggtg ccattcgcag 1020
gcgcctgctg ctcctcttcc tgggaatcgc cgagggaaag cctcgacttg cccgctgttt 1080
ccatcgattt ccaggtgcct ccactggtgg cctgctgcag cgaaatgcga ggtatggact 1140
tgcagctgcc caagatagag gtgctatctc cgggacggtc ctcggattcg gccagggcgc 1200
tgagactgca agtgggcggc aggatgcggt gtggaaagcg tggcgaaccc cttacccttt 1260
cccttaaact gggtggctgc tgcaccaccg aaggtctccg gttgctgggc aattggaagt 1320
tcaccttggt ctccagcagt ttcacatact ttccacccaa cgtttggtcc ccctgttcac 1380
cgccagaaat agtggggata ggcaggaaat tcggatcctt ggtaaggtga tccaggtgag 1440
taccgctttt ggtcaatgaa cccctatgga tgggctttag gtggttacca ttcgggtggg 1500
aggacgcctc tgagggaatg gaatcctgat ttggattatt gtatatggca ttgcaggacg 1560
ccgtttgaga ggagggaatc aggcgggtga acagtcccat gtgcgattta aaggccgcct 1620

tcagaccctt ttggtccgcc aagctggcgg tggatccctt gttgctcgaa gtggtgcgag 1680
aaaatatccc tcgtaatcgg tttatcaagc ccgtcttagc ctccgtggcc tgctggtggg 1740
ctaaacggag gttaatcaag tcctgtgcag gcaggcaggc ctcctctaat tggggctggg 1800
aggcaaaggt aacagaacgc ctcatcggga tctgcatgta ctctgctgac ggccaactgg 1860
cccgtggaat cacaggaggc agcacggaaa gcaggagatt gggcagtatt cccgagatac 1920
ttggagttct gtttttactt gcagccggag tctgcgagat gccactggtt accgtcagtt 1980
ctggctttag atgacctccc ccagttaaga gtgatgtgga cgacgaagtt gtggtggagg 2040
tggtggttag tcccgactcg agtagcctct ccagctttcg aatctcggcg tccagagcct 2100
ggatttcctt cttatagacc ctgttttgca cctccactcg atactgcagt tctcggcgat 2160
catcgaccac gaatcggcgt gtgttgcact ccatcttaag gcccatactg tggataaccg 2220
gatcgataat atcgtctgag ggagaaggag agttcccaaa ggattaggac acaactttat 2280
aaaggttcta gaaattagga gactagaact taaaagtatt tcctaagtga gcaagatctt 2340
attgcacctt caagcgctac ttacttcttg cccaaatatc atggagtttt gggatgaaaa 2400
gcagacaaag ggttgcagtg gtgctggtta aaatcagagc tgtgattgtg atgaaggcca 2460
gggtgactcg ctccgaaatc aagttggcca gcaccacgac gatggcgctg gtgatgacca 2520
cactgtatac agacactccg atgtactgcg agtcattgag agcaggtatt tttacgtggc 2580
gcgtctccca ggccatatag acacccacca caagaagaag gcctttgtag gcgtacagga 2640
cactcaacca cgtttgcgtg tgctgcgaac ggcaaacttc aacctaaaat agatagatag 2700
atagatattc atagattcta aatgaattaa tactttttca acataattaa aagtaatctg 2760
ataatttgaa gcatattaat atttttctta cctgaggctg gtaaacgaca cttctatcag 2820
tcgcactgat ctcgagcgtc aggttgtgaa gatggcgctc cattggatcg gtgaccaccc 2880
aaagggttac gagcagcgca tccaccagaa gcaatccgcc gacgagcaag atcagttgaa 2940
tgtcctgcag catcttgtcc ttgaaaacgc tgccggtacg agtgaatatc cgatgcactc 3000
tgtaggtctt ggcaaacatc gatccaaagg ccaacgagaa tccggcggag agcagataga 3060
cgcgggcctg gtgagcatga taagcggatg acggggaagt tcaggttagg gagaaagtac 3120
tgtagtagca gcagcaaagt aacacactaa gtgtcatgtc aggcagaccc aaagaaaaca 3180
agccggcggc aacaatgggg gaaacaagtc atcaatggac cacaattctc agagtgcttg 3240
agactttgtt gtgcgtttca aatttatgca aataaaggaa tatcacaaaa gttgtaaatt 3300
tccgtgtcat gtgtttgcat ttccctgcca caagcctcaa caaaacaaaa acaagcaggc 3360
aactggaaaa gcggaacgac agtagaagtg gttacagaaa cagaaacaac agcgaccagc 3420
aaatttccca ggagtggcgt tcgaggcagt tgttaagggg gtcaaagtgg aaaggcgcag 3480
atcgggatag aagaaagctc gtcaaacttg acttgacttc tggcttaaag agcagcagat 3540
ttgcagcgaa accgaacggc tatttatagc cagcaggggc agtcgtgtat cttgaaaaga 3600
cacgcggaat aatcaatatt cagtttatat aatccacttt aaaataatag aaattaacaa 3660
aaaattgcca caattatcag taatttttat acttcgcagg ctattttatc gccatcttcg 3720
tgtatacgaa tacgacaatt agttcaagca atatttcttt cagtgaataa ataaagcaaa 3780
cacatgagcg cgactcgaaa gttaccagcc tgaaactgaa agtcatattt gatgaggacc 3840
ccacatgctt gcttacacac agaagcccgt caaagtcaaa caaagacggc tgagaggaaa 3900
ggcggatgag aatgcggatt cggataggaa tggcgatagg gatggggatg tcggagccga 3960
aggataacaa ctcaccgtgc agaccgttgc gaaagagtcc tccgccgagg gcagcgtcga 4020
gtggtccaag cccaaaagga tgacggtggc gtacacaaag atgcagccca ctgcggtgat 4080
gttgctcagc ttcgggctgg aaagtttaat tgccctgtgg agtgagaagg gggatggtca 4140
tggttataag ctgaggcgcc gcaatccttt gcagagaaag tgcagagtaa ccgtagaatg 4200
cgaaacatat atacgaactc aactaacgat gggtcagaaa tacagctgga aagccaagca 4260
attacttcac ttcaaaaacc gaaacaagtc acatgtgtaa tgggttcaat gggttgagca 4320
ccgaaagcag ttaactcatc atggaatata aagaaacagc ccgcacaaat tctcttttat 4380
atggccacta aatattttat agtcgtttaa tgaaaaataa tatattcaaa caatcagggc 4440
ccataaaaag agtaagccaa taatttatgt acttatcgcc ttacttcagc ttccgaaagt 4500
gcagattgaa cgccaggaag gcgatggcca gagcgattcc cacgctggag agggtggcga 4560
tggtgtagaa ggccagtgga gcgatggtcg ccacccgcag cttgaacacc cgcttggcga 4620
tgggtacctg cccgctgtgc cacttcaccg gccggcagcg gggacaccgg aagtccaggg 4680
catccgtggc cggatagtag agggccaccg gttccagcaa accgcgctgg atttgataga 4740
aggcagtggt gccaacgcga tctgggccgc tgaaggaaac ggggccctac caaaaagaga 4800
aatactcgtg atgaactaga tttaagaacc ttaggatcaa agttaaatgg gttggatgat 4860
ttcagttgaa agagataact caccgacact cccaggaagt ggagcttgcc catttgctgc 4920
aggaactccc aggccatgtc gctgcgggtg taatcgaatc cgtccagctt cgactgctcc 4980
tcgttccgcc gccagtgctc ctcagcggct ctcaaggcca gggcgatggc ccacacggcg 5040
tcgtaggttt gcggcgcgta ctgcgagatc gcctccggaa agaggtgccc tccgctggtg 5100
cctcccacgc cccttctcct ccgccgacga gagtcagatt gcgttgcagc gatactgatc 5160
ctggaaccat aaccggagcc aaatccatcc tggccgtgga actgggcgga ttgcttgcgc 5220

agctgggagt tgaacatgtg attgttctgc gaagtggaaa cgcgattaaa caagcctcga 5280
atgccgcaga aatgtcgcca gtaatgagat tgcctatggc agctaccact cctgccccaa 5340
tttcatgacc agcactaaaa gagccataaa aacgcgtctg cgatggtgaa tgcgaatgcg 5400
atgacaggcc ggacttaggg ggcttcgcac ttggcaatca ccatcgccat cgccattgct 5460
gaaaccatcg cgaaacgcaa tcgcagcctc gggattgtct aatcaaataa atcatttcat 5520
ttcggacacg cacagcaaga cggacaagac aggacaacga gaaggacaag atagtcgtct 5580
ctgaataaaa aaaaaaaaaa aaaaaaacgg ggaataaaag tttcggcaat tctggaaaat 5640
gttgcttgaa gctgcctttg agagttgggt tcaatccaat ttggatcccg aaatttatat 5700
gcatgtataa cttaacaaca atctattctt gaaattcatc tttacgttaa tctcatactg 5760
tcatgaaatt aatttattta ttttcccttt aatattgaat gatattttct cgcagtgcac 5820
ctcctcaact taccagtcca ctatagctga cgttatttcc aacgatgctg ttgtgcgttg 5880
agaccacgat gaggttctcg acggccagct gcagttcgtg gttagagcag gcggtgcgct 5940
ggtccggcca ccacggagcc cccatgctct cgtggaggat ccaggcgtag tccgccccga 6000
acattcgaag cctgtaggcc tcgcacagga tctgggggc cagctcctgc gagaagctgc 6060
cgatgatgat gcgcgtgtcc gtctcctgtg gcgatacgaa cgaaacggaa cggcatgtgc 6120
ccatttgtca tttgtcatcc ttcgattagc acaagtgcca atccacactc aaaaaaaatg 6180
ggtaacaatc atcatgctga tagagggtct ctaggcgatg tggtctgcta atgaaacatt 6240
tatccatgcg gactggacga acaaaaaact tgaagataaa ctcagagaaa taagttaacg 6300
taaaaaacgt aattataatc atacatatga acctttttt taaaaactta ctaaattaag 6360
ttgccaggca aataaattgg tctatttaaa tattccccaa atacgttcat tttcatattt 6420
gttcatgcat aagcttttaa tataaacaga acagcgctag caccatttta atttaaacat 6480
tttcgaaaat tgaaatttag ttaaatttaa tgttggtttt tcaaacaaat attcataaat 6540
gtttacgttt atacgtttgc gtatacgagt ataaaataaa acgaaaatgt gtatttaaat 6600
ggccatttgt tgctcaccct aagtagcagc agctgctcct tgaagtcggt ggccgcaaag 6660
gtgatggtgg cggcacagga tatgttggcc gcctccagtt cggtgaccag gttgttcacc 6720
gccagcgagt gaacctcctc gttctgcgag aaagtggtca ccgtgcccca gccaaacttc 6780
cgaatgaaag cgatgcgcgc cggattgtgt gaggagtccg gggccaccgt cctgtagaag 6840
taggggaact cccgcctgtc gctcaacgcc ggcgatgtgg aaccgaagga tacctttcat 6900
cgaacgaaat cgttagacta tctctaccgt acatggggtc gcttaaaatg caagacacgc 6960
atcaagaaat ccagccctct aggtcatgtg tcttcccagt ttccccactc atttaacaca 7020
cgcaccgccg aaatgaaact tcaaaggcca aatgcaaatg ctctaagtga atttgtggac 7080
ctcataattt gtgggctgtg gaaatgttta tggctcaaag gtttatgggg tgctggaatc 7140
ggctggattc cgcagcattt atgctaatcg agctacgcac gtgagttgat tagattttg 7200
agaaattgcc cgatacgagg agttggttgg atgaaccgag cgagctggat tatttgcatt 7260
ttaatcagtc gcagatgact cacctgcacg atgttccagt agggcaccac cttcgccagg 7320
ctctcggtga cctccgagca ggccgatccc agcagcatca ccatcctcgt cgagggctgt 7380
gtgtagatgg cgtggaagaa gcgatccacg cccactccag gatcacactg attgtcggag 7440
aagaaatgtg tgaagaaacg gaggagcaaa ggacaccaca tcagagatga gtaatagagg 7500
gggcagggaa aatagcagtg gcagaggccg ttgctcgctg actcgtttcg acgcaataaa 7560
aaggcaacat gatggcgcaa taaacatttt taaaatgtcg acttttagag aggcacagcc 7620
ggagatgcgg ctccagatat gacagtctgt ttaaggtcga tgtgcatatg catacactta 7680
aaaaaatcgt ttgaaaacct agttcactaa gtgctgacat cacagtcgag gtattacgta 7740
ttgacggttt aatttatcag aaactaaact actattatgc attttctttt ttttaccatc 7800
aattaattgt agctttaata cactactgca ccaagttaaa aatgattgca ggtaaatagt 7860
ttacattgtt ggtaacttag tcagtggcat tgcttttctt cagtgacagt gttctgtaca 7920
tttgcgtgcc ccatcaaaac cggcaacgac tcatcgcaat aaaagcgcga aagacggaag 7980
gttgatgaca tatttgcctc tcttttaggg ccacagccgc agctgctatc cgcactgttc 8040
ttgttgttgc agattttctt attaagtttt aatgttgcac gcgctactga atgaaaattc 8100
aatttcgtaa tggaacagcg acgcctagaa gtggcgccca acgtggggca tacggttacg 8160
ccacgtagaa atgacattct cacctgagta tcgttggtca cgagctcgag ggtgtagccc 8220
ggcagcaggc gcttgcggtt gatgtgctcc acggccatgg tggcagctcc caattcgctc 8280
agaccatccg gacgtggtcc ccgcgatgtg gacagctcaa agagtccgag cagcacgatt 8340
ttgccatgat tgcgctgcat ttcggccggc gacatgcggc gttcgctttc gaacacctgg 8400
ttgaccttca agtatcgctg cctgggggc agattgagga ggtggtgggc gtgggaggca 8460
ttggtgcttc cgggcagatt gctgtgcgaa ttgaggcgct gctgccggcg tttgactttc 8520
gtccgtttgt acttccaatc ggagcgcggt ggaatcggag tccccctgc agtcacagtc 8580
gatgccgatt ctccgggtgg cgagctggat gggctggatg tgctaattga ggcagcacct 8640
ggcctaattg cctcgtagcc cagggcctgc agctcggctg atgactccat ggcagccgct 8700
gccgacgttg cccaggcaag ggccaggact agcctcagtc ccacggccgg ccatggaccg 8760
tatctggtcc cttggaccgg ttgaattatg cgcatcccgt cggcaatggc caatggccca 8820

ttaatcacca aggcatttga atcgtcggct atcagaggct gcagcaacaa acaaacacga 8880
gccacacagc acaggacacc ggaatatata tcgccttttt atgcactctg ccgacaggat 8940
acgcttatcg tttctcttcc gacgagcggc tcttccttct gctccctcct cttgtatcgg 9000
cgtcaatttt atcgcctttt tattggcctc actttgctc gcattttgta tctttttgca 9060
cattcactct gcacagcagt cgaccataaa agtatgcaat gcaaaggact ttctcgatag 9120
gtccgattga caggattggg taacacggga aaaaactaca cgccattcgc taactttcag 9180
tcacttttcg tagaaaaaaa tgcttatttt tattatgcac tgtttattaa ggattaacac 9240
aatttaggtg aatatttatt agtattttgc tgtagtatgt gattttatta aaatgtaatt 9300
tgaatctaca tatgtatata tatgtatata tgtgaaaaca ttgatttttt catgagaatg 9360
tgaaagggaa tgttaatgga aataccgaaa caaatgtcta gggaggcata aaaaggcatt 9420
tcacaccaat ttggcagtcg aaaggattat gacgaaggct tcgatgatat ggacacaccg 9480
cgccgctcac gtctcgtccc atgacttgtt atgtgcctgc cacttggcca ttattttta 9540
tgctaatggc caaacatttt gcgctttcat tgcgggcttt tggactactt tttgtcgagt 9600
ctcacgcgcc acagatttct gtggcgaacg tttttctctg tatgcgcgcg tatgcataat 9660
tgtttcaaga ttaccaactt gcagtccctt tgccgctttc ccgcttttcc accgcccact 9720
tgcctgcacc taagcccact tcatagacgc tttccttcgc gctttctgct tttcccacag 9780
cacgcttttc gctgt 9795

<210> 14
<211> 3918
<212> DNA
<213> Drosophila

<400> 14
atgcgcataa ttcaaccggt ccaagggacc agatacggtc catggccggc cgtgggactg 60
aggctagtcc tggcccttgc ctgggcaacg tcggcagcgg ctgccatgga gtcatcagcc 120
gagctgcagg ccctgggcta cgaggcaatt aggccaggtg ctgcctcaat tagcacatcc 180
agcccatcca gctcgccacc cggagaatcg gcatcgactg tgactgcagg ggggactccg 240
attccaccgc gctccgattg gaagtacaaa cggacgaaag tcaaacgccg gcagcagcgc 300
ctcaattcgc acagcaatct gcccggaagc accaatgcct cccacgccca ccacctcctc 360
aatctgcccc ccaggcagcg atacttgaag gtcaaccagg tgttcgaaag cgaacgccgc 420
atgtcgccgg ccgaaatgca gcgcaatcat ggcaaaatcg tgctgctcgg actctttgag 480
ctgtccacat cgcggggacc acgtccggat ggtctgagcg aattgggagc tgccaccatg 540
gccgtggagc acatcaaccg caagcgcctg ctgccgggct acaccctcga gctcgtgacc 600
aacgatactc agtgtgatcc tggagtgggc gtggatcgct tcttccacgc catctacaca 660
cagccctcga cgaggatggt gatgctgctg ggatcggcct gctcggaggt caccgagagc 720
ctggcgaagg tggtgcccta ctggaacatc gtgcaggtat ccttcggttc cacatcgccg 780
gcgttgagcg acaggcggga gttcccctac ttctacagga cggtggcccc ggactcctca 840
cacaatccgg cgcgcatcgc tttcattcgg aagtttggct ggggcacggt gaccactttc 900
tcgcagaacg aggaggttca ctcgctggcg gtgaacaacc tggtcaccga actggaggcg 960
gccaacatat cctgtgccgc caccatcacc tttgcggcca ccgacttcaa ggagcagctg 1020
ctgctactta gggagacgga cacgcgcatc atcatcggca gcttctcgca ggagctggcc 1080
ccccagatcc tgtgcgaggc ctacaggctt cgaatgttcg ggcggacta cgcctggatc 1140
ctccacgaga gcatgggggc tccgtggtgg ccggaccagc gcaccgcctg ctctaaccac 1200
gaactgcagc tggccgtcga gaacctcatc gtggtctcaa cgcacaacag catcgttgga 1260
aataacgtca gctatagtgg actgaacaat cacatgttca actcccagct gcgcaagcaa 1320
tccgcccagt tccacggcca ggatggattt ggctccggtt atggttccag gatcagtatc 1380
gctgcaacgc aatctgactc tcgtcggcgg aggagaaggg gcgtgggagg caccagcgga 1440
gggcacctct ttccggaggc gatctcgcag tacgcgccgc aaacctacga cgccgtgtgg 1500
gccatcgccc tggccttgag agccgctgag gagcactggc ggcggaacga ggagcagtcg 1560
aagctggacg gattcgatta cacccgcagc gacatggcct gggagttcct gcagcaaatg 1620
ggcaagctcc acttcctggg agtgtcgggc cccgtttcct tcagcggccc agatcgcgtt 1680
ggcaccactg ccttctatca aatccagcgc ggtttgctgg aaccggtggc cctctactat 1740
ccggccacgg atgccctgga cttccggtgt ccccgctgcc ggccggtgaa gtggcacagc 1800
gggcaggtac ccatcgccaa gcgggtgttc aagctgcggg tggcgaccat cgctccactg 1860
gccttctaca ccatcgccac cctctccagc gtgggaatcg ctctggccat cgccttcctg 1920
gcgttcaatc tgcactttcg gaagctgaag gcaattaaac tttccagccc gaagctgagc 1980
aacatcaccg cagtgggctg catctttgtg tacgccaccg tcatcctttt gggcttggac 2040

```
cactcgacgc tgccctcggc ggaggactct ttcgcaacgg tctgcacggc ccgcgtctat 2100
ctgctctccg ccggattctc gttggccttt ggatcgatgt ttgccaagac ctacagagtg 2160
catcggatat tcactcgtac cggcagcgtt ttcaaggaca agatgctgca ggacattcaa 2220
ctgatcttgc tcgtcggcgg attgcttctg gtggatgcgc tgctcgtaac cctttgggtg 2280
gtcaccgatc caatggagcg ccatcttcac aacctgacgc tcgagatcag tgcgactgat 2340
agaagtgtcg tttaccagcc tcaggttgaa gtttgccgtt cgcagcacac gcaaacgtgg 2400
ttgagtgtcc tgtacgccta caaaggcctt cttcttgtgg tgggtgtcta tatggcctgg 2460
gagacgcgcc acgtaaaaat acctgctctc aatgactcgc agtacatcgg agtgtctgta 2520
tacagtgtgg tcatcaccag cgccatcgtc gtggtgctgg ccaacttgat ttcggagcga 2580
gtcaccctgg ccttcatcac aatcacagct ctgattttaa ccagcaccac tgcaaccctt 2640
tgtctgcttt tcatcccaaa actccatgat atttgggcaa gaaacgatat tatcgatccg 2700
gttatccaca gtatgggcct taagatggag tgcaacacac gccgattcgt ggtcgatgat 2760
cgccgagaac tgcagtatcg agtggaggtg caaaacaggg tctataagaa ggaaatccag 2820
gctctggacg ccgagattcg aaagctggag aggctactcg agtcgggact aaccaccacc 2880
tccaccacaa cttcgtcgtc cacatcactc ttaactgggg gaggtcatct aaagccagaa 2940
ctgacggtaa ccagtggcat ctcgcagact ccggctgcaa gtaaaaacag aactccaagt 3000
atctcgggaa tactgcccaa tctcctgctt tccgtgctgc ctcctgtgat tccacgggcc 3060
agttggccgt cagcagagta catgcagatc ccgatgaggc gttctgttac ctttgcctcc 3120
cagccccaat tagaggaggc ctgcctgcct gcacaggact tgattaacct ccgtttagcc 3180
caccagcagg ccacggaggc taagacgggc ttgataaacc gattacgagg gatatttcct 3240
cgcaccactt cgagcaacaa gggatccacc gccagcttgg cggaccaaaa gggtctgaag 3300
gcggccttta aatcgcacat gggactgttc acccgcctga ttccctcctc tcaaacggcg 3360
tcctgcaatg ccatatacaa taatccaaat caggattcca ttccctcaga ggcgtcctcc 3420
caccccgaatg gtaaccacct aaagcccatc catagggggtt cattgaccaa aagcggtact 3480
cacctggatc accttaccaa ggatccgaat ttcctgccta tccccactat ttctggcggt 3540
gaacaggggg accaaacgtt gggtggaaag tatgtgaaac tgctggagac caaggtgaac 3600
ttccaattgc ccagcaaccg gagaccttcg gtggtgcagc agccacccag tttaagggaa 3660
agggtaaggg gttcgccacg ctttccacac cgcatcctgc cgcccacttg cagtctcagc 3720
gccctggccg aatccgagga ccgtcccgga gatagcacct ctatcttggg cagctgcaag 3780
tccatacctc gcatttcgct gcagcaggcc accagtggag gcacctggaa atcgatggaa 3840
acagcgggca agtcgaggct ttccctcggc gattcccagg aagaggagca gcaggcgcct 3900
gcgaatggca ccgaataa                                                3918
```

<210> 15
<211> 1305
<212> PRT
<213> Drosophila

<400> 15
```
Met Arg Ile Ile Gln Pro Val Gln Gly Thr Arg Tyr Gly Pro Trp Pro
 1               5                  10                  15
Ala Val Gly Leu Arg Leu Val Leu Ala Leu Ala Trp Ala Thr Ser Ala
            20                  25                  30
Ala Ala Ala Met Glu Ser Ser Ala Glu Leu Gln Ala Leu Gly Tyr Glu
        35                  40                  45
Ala Ile Arg Pro Gly Ala Ala Ser Ile Ser Thr Ser Ser Pro Ser Ser
    50                  55                  60
Ser Pro Pro Gly Glu Ser Ala Ser Thr Val Thr Ala Gly Gly Thr Pro
65                  70                  75                  80
Ile Pro Pro Arg Ser Asp Trp Lys Tyr Lys Arg Thr Lys Val Lys Arg
                85                  90                  95
Arg Gln Gln Arg Leu Asn Ser His Ser Asn Leu Pro Gly Ser Thr Asn
            100                 105                 110
Ala Ser His Ala His His Leu Leu Asn Leu Pro Pro Arg Gln Arg Tyr
        115                 120                 125
Leu Lys Val Asn Gln Val Phe Glu Ser Glu Arg Arg Met Ser Pro Ala
    130                 135                 140
Glu Met Gln Arg Asn His Gly Lys Ile Val Leu Leu Gly Leu Phe Glu
145                 150                 155                 160
Leu Ser Thr Ser Arg Gly Pro Arg Pro Asp Gly Leu Ser Glu Leu Gly
```

```
                165                 170                 175
Ala Ala Thr Met Ala Val Glu His Ile Asn Arg Lys Arg Leu Leu Pro
            180                 185                 190
Gly Tyr Thr Leu Glu Leu Val Thr Asn Asp Thr Gln Cys Asp Pro Gly
        195                 200                 205
Val Gly Val Asp Arg Phe Phe His Ala Ile Tyr Thr Gln Pro Ser Thr
    210                 215                 220
Arg Met Val Met Leu Leu Gly Ser Ala Cys Ser Glu Val Thr Glu Ser
225                 230                 235                 240
Leu Ala Lys Val Val Pro Tyr Trp Asn Ile Val Gln Val Ser Phe Gly
                245                 250                 255
Ser Thr Ser Pro Ala Leu Ser Asp Arg Arg Glu Phe Pro Tyr Phe Tyr
            260                 265                 270
Arg Thr Val Ala Pro Asp Ser Ser His Asn Pro Ala Arg Ile Ala Phe
        275                 280                 285
Ile Arg Lys Phe Gly Trp Gly Thr Val Thr Thr Phe Ser Gln Asn Glu
    290                 295                 300
Glu Val His Ser Leu Ala Val Asn Asn Leu Val Thr Glu Leu Glu Ala
305                 310                 315                 320
Ala Asn Ile Ser Cys Ala Ala Thr Ile Thr Phe Ala Ala Thr Asp Phe
                325                 330                 335
Lys Glu Gln Leu Leu Leu Leu Arg Glu Thr Asp Thr Arg Ile Ile Ile
            340                 345                 350
Gly Ser Phe Ser Gln Glu Leu Ala Pro Gln Ile Leu Cys Glu Ala Tyr
        355                 360                 365
Arg Leu Arg Met Phe Gly Ala Asp Tyr Ala Trp Ile Leu His Glu Ser
    370                 375                 380
Met Gly Ala Pro Trp Trp Pro Asp Gln Arg Thr Ala Cys Ser Asn His
385                 390                 395                 400
Glu Leu Gln Leu Ala Val Glu Asn Leu Ile Val Val Ser Thr His Asn
                405                 410                 415
Ser Ile Val Gly Asn Asn Val Ser Tyr Ser Gly Leu Asn Asn His Met
            420                 425                 430
Phe Asn Ser Gln Leu Arg Lys Gln Ser Ala Gln Phe His Gly Gln Asp
        435                 440                 445
Gly Phe Gly Ser Gly Tyr Gly Ser Arg Ile Ser Ile Ala Ala Thr Gln
    450                 455                 460
Ser Asp Ser Arg Arg Arg Arg Arg Arg Gly Val Gly Gly Thr Ser Gly
465                 470                 475                 480
Gly His Leu Phe Pro Glu Ala Ile Ser Gln Tyr Ala Pro Gln Thr Tyr
                485                 490                 495
Asp Ala Val Trp Ala Ile Ala Leu Ala Leu Arg Ala Ala Glu Glu His
            500                 505                 510
Trp Arg Arg Asn Glu Glu Gln Ser Lys Leu Asp Gly Phe Asp Tyr Thr
        515                 520                 525
Arg Ser Asp Met Ala Trp Glu Phe Leu Gln Gln Met Gly Lys Leu His
    530                 535                 540
Phe Leu Gly Val Ser Gly Pro Val Ser Phe Ser Gly Pro Asp Arg Val
545                 550                 555                 560
Gly Thr Thr Ala Phe Tyr Gln Ile Gln Arg Gly Leu Leu Glu Pro Val
                565                 570                 575
Ala Leu Tyr Tyr Pro Ala Thr Asp Ala Leu Asp Phe Arg Cys Pro Arg
            580                 585                 590
Cys Arg Pro Val Lys Trp His Ser Gly Gln Val Pro Ile Ala Lys Arg
        595                 600                 605
Val Phe Lys Leu Arg Val Ala Thr Ile Ala Pro Leu Ala Phe Tyr Thr
    610                 615                 620
Ile Ala Thr Leu Ser Ser Val Gly Ile Ala Leu Ala Ile Ala Phe Leu
625                 630                 635                 640
Ala Phe Asn Leu His Phe Arg Lys Leu Lys Ala Ile Lys Leu Ser Ser
```

```
                645                 650                 655
Pro Lys Leu Ser Asn Ile Thr Ala Val Gly Cys Ile Phe Val Tyr Ala
            660                 665                 670
Thr Val Ile Leu Leu Gly Leu Asp His Ser Thr Leu Pro Ser Ala Glu
        675                 680                 685
Asp Ser Phe Ala Thr Val Cys Thr Ala Arg Val Tyr Leu Leu Ser Ala
    690                 695                 700
Gly Phe Ser Leu Ala Phe Gly Ser Met Phe Ala Lys Thr Tyr Arg Val
705                 710                 715                 720
His Arg Ile Phe Thr Arg Thr Gly Ser Val Phe Lys Asp Lys Met Leu
                725                 730                 735
Gln Asp Ile Gln Leu Ile Leu Leu Val Gly Gly Leu Leu Leu Val Asp
            740                 745                 750
Ala Leu Leu Val Thr Leu Trp Val Val Thr Asp Pro Met Glu Arg His
        755                 760                 765
Leu His Asn Leu Thr Leu Glu Ile Ser Ala Thr Asp Arg Ser Val Val
    770                 775                 780
Tyr Gln Pro Gln Val Glu Val Cys Arg Ser Gln His Thr Gln Thr Trp
785                 790                 795                 800
Leu Ser Val Leu Tyr Ala Tyr Lys Gly Leu Leu Leu Val Val Gly Val
                805                 810                 815
Tyr Met Ala Trp Glu Thr Arg His Val Lys Ile Pro Ala Leu Asn Asp
            820                 825                 830
Ser Gln Tyr Ile Gly Val Ser Val Tyr Ser Val Val Ile Thr Ser Ala
        835                 840                 845
Ile Val Val Val Leu Ala Asn Leu Ile Ser Glu Arg Val Thr Leu Ala
    850                 855                 860
Phe Ile Thr Ile Thr Ala Leu Ile Leu Thr Ser Thr Thr Ala Thr Leu
865                 870                 875                 880
Cys Leu Leu Phe Ile Pro Lys Leu His Asp Ile Trp Ala Arg Asn Asp
                885                 890                 895
Ile Ile Asp Pro Val Ile His Ser Met Gly Leu Lys Met Glu Cys Asn
            900                 905                 910
Thr Arg Arg Phe Val Val Asp Asp Arg Arg Glu Leu Gln Tyr Arg Val
        915                 920                 925
Glu Val Gln Asn Arg Val Tyr Lys Lys Glu Ile Gln Ala Leu Asp Ala
    930                 935                 940
Glu Ile Arg Lys Leu Glu Arg Leu Leu Glu Ser Gly Leu Thr Thr Thr
945                 950                 955                 960
Ser Thr Thr Thr Ser Ser Ser Thr Ser Leu Leu Thr Gly Gly Gly His
                965                 970                 975
Leu Lys Pro Glu Leu Thr Val Thr Ser Gly Ile Ser Gln Thr Pro Ala
            980                 985                 990
Ala Ser Lys Asn Arg Thr Pro Ser Ile Ser Gly Ile Leu Pro Asn Leu
        995                 1000                1005
Leu Leu Ser Val Leu Pro Pro Val Ile Pro Arg Ala Ser Trp Pro Ser
    1010                1015                1020
Ala Glu Tyr Met Gln Ile Pro Met Arg Arg Ser Val Thr Phe Ala Ser
1025                1030                1035                1040
Gln Pro Gln Leu Glu Glu Ala Cys Leu Pro Ala Gln Asp Leu Ile Asn
                1045                1050                1055
Leu Arg Leu Ala His Gln Gln Ala Thr Glu Ala Lys Thr Gly Leu Ile
            1060                1065                1070
Asn Arg Leu Arg Gly Ile Phe Ser Arg Thr Thr Ser Ser Asn Lys Gly
        1075                1080                1085
Ser Thr Ala Ser Leu Ala Asp Gln Lys Gly Leu Lys Ala Ala Phe Lys
    1090                1095                1100
Ser His Met Gly Leu Phe Thr Arg Leu Ile Pro Ser Ser Gln Thr Ala
1105                1110                1115                1120
Ser Cys Asn Ala Ile Tyr Asn Asn Pro Asn Gln Asp Ser Ile Pro Ser
```

```
                1125                      1130                      1135
Glu Ala Ser Ser His Pro Asn Gly Asn His Leu Lys Pro Ile His Arg
            1140                      1145                      1150
Gly Ser Leu Thr Lys Ser Gly Thr His Leu Asp His Leu Thr Lys Asp
        1155                      1160                      1165
Pro Asn Phe Leu Pro Ile Pro Thr Ile Ser Gly Gly Glu Gln Gly Asp
    1170                      1175                      1180
Gln Thr Leu Gly Gly Lys Tyr Val Lys Leu Leu Glu Thr Lys Val Asn
1185                      1190                      1195                      1200
Phe Gln Leu Pro Ser Asn Arg Arg Pro Ser Val Val Gln Gln Pro Pro
                1205                      1210                      1215
Ser Leu Arg Glu Arg Val Arg Gly Ser Pro Arg Phe Pro His Arg Ile
            1220                      1225                      1230
Leu Pro Pro Thr Cys Ser Leu Ser Ala Leu Ala Glu Ser Glu Asp Arg
        1235                      1240                      1245
Pro Gly Asp Ser Thr Ser Ile Leu Gly Ser Cys Lys Ser Ile Pro Arg
    1250                      1255                      1260
Ile Ser Leu Gln Gln Ala Thr Ser Gly Gly Thr Trp Lys Ser Met Glu
1265                      1270                      1275                      1280
Thr Ala Gly Lys Ser Arg Leu Ser Leu Gly Asp Ser Gln Glu Glu Glu
                1285                      1290                      1295
Gln Gln Ala Pro Ala Asn Gly Thr Glu
            1300                      1305
```

<210> 16
<211> 7913
<212> DNA
<213> Drosophila

<400> 16

```
ttttctgccc gtttgtttgt gctcccgtcc aggtgaaggt taagtgctcc gagctgagga 60
tcaaggacaa gaaggaactc accaagcaat tggatgagct caagaatgag ttgctcagcc 120
tgcgcgtggc caaggtgacc ggcggagctc cctccaagct ctccaagatg tgagtatgcc 180
gcggcacaac gacaagcgtc acgaactaat tccggtgttt tgtgcattta tgtgcgttct 240
tatccataca gccgcgttgt ccgcaaggcc atcgctcgcg tctacattgt gatgcaccag 300
aagcagaagg agaatctgcg caaggtcttc aagaacaaga agtacaagcc cctggatctg 360
cgcaagaaga agacccgcgc tatccgcaag gccctgtctc cgcgcgacgc caaccgcaag 420
accctcaagg agatccgcaa gcgctccgtc ttcccccaga ggaagttcgc cgtcaaggcc 480
tagaatgcat ccagatactc cgactgtcta tagtgcgcgt taagggttct tcggtacaat 540
gaagaagaaa aactaaataa aatctttttt ttaaagtgaa aatagcgatg cgtgtttgat 600
tttgtgaccg aaaaaacgcg agtggcacat ggctttggcc atgggttcat ggaaattagt 660
taaaaaacaa aaaaacaagt gccgcattcc ctgttctaca tatgttagtc ggatccaatg 720
ggatctactt aactacttgt ggtccttaag atcattatag agcaaaaacg ttgagttagc 780
gccttgtaga agtaatggca aatcttaagt tcaatcttta attcagcctt ctttgaatat 840
gcaaacttga aaaacaccag aattctcaat tggtctacgt atatgtattt aattcgattg 900
tctttgatgt aagcacggcc atacatcgta aataattatg aactaacaag taataagtag 960
gtaattgata tgaattacaa gctctcgact gcgagctcga tcatactgaa tatgtttggc 1020
acatttggtt aatactattc ttctgcacag cgtgaaacag cggcgtgctg cggatgacgc 1080
ccgacagctc cttggagttc ttgctttggt gatagtttct actcggcatc ggcgacaacg 1140
gttcgcccca gaacttgagc agcgcaaaga tcttgtagta ggccactctg aaatagaaga 1200
tagattcgct attagcagct aaacccgata gcagccaaaa ctgccagcaa acttacctgt 1260
agttgcgatt gctggccgca tagatgatcg ggttaatgac actggaagcc caggccatca 1320
cggaggcgat gatgtgcagc caggggtacg aggtattgcg ctcgtcgtcc accacattgg 1380
ccagcattag cggcaggaag cagacgagga agcacagaaa gatggtgacc atcatgacgg 1440
tcagtcgatt gtcttcgcgc gccttgcgcg tgcatgtggt ggtcatgtag gatccgccgc 1500
cggacgacga ggagcccttg gccgctgcga tctggaagtt atcgtggttg cgaatcttct 1560
ttttctggtg cagcaccgta atgtagatgc acgagtacga aacgatgatc accaggcaag 1620
gcagcaggaa gccgatcacg aacagcgtct tcttgatcga tcgcccctcc ttcttgagta 1680
ttgtgcagga gaaggtggcc tcgtccaggc ccatctcgcc ccagatgccc aatatgggcg 1740
```

gcagctggga agataaaaac cgattagttg aatgtgtgca cacaaatcga agacgcctta 1800
cttaccaaca gcagaaagga gacggcccaa acgaacaaca gctgcagggt tatgaactta 1860
ggcttatata tctgcgagta gcggctgtgg caagcgatga gtatatatct gccaggggat 1920
tagggtgacg gttgcggatt agcttgcgac aatgcggact agacttgggc ctacctgttc 1980
agggtgatgc ccaccatgct gaggagtgaa acagccacat tgccatagaa gatcaccgga 2040
aagatcttgc acaatgtggt accaaaagtc cagctctgtc gttggaatcg agcaaattat 2100
aatcattcac taattgtata ttacattcac aggtataatg actcacctcc tggaagaatc 2160
gcactgcggt cagtggcagg ctgaaggagc agaagagcag gtcggagatg cttagcgaaa 2220
tgacgaaggc ggttgtggca tgctcccgta tcgtggggct cttgagcagc gccagcaggg 2280
taatcaggtt gcctgtgcgg caattcagca tagagaaaca ggttagctaa tcgaggagct 2340
gaatgtggga tgccaattac tcaccaagaa cgccgatcgt cacaaagaca caggcactaa 2400
tggccgcgaa taatgtcgcc gaatgcgggt agatcgattg cgtggcggcc gccggttcat 2460
ccatctgcat gtctgcgtcc tgaaagtagc ccgttgccat gcccatgtcc tgatccattc 2520
tgcgggggaa acaagcaaac gatgtgatta gccaatttca attacttatt tttgaacagg 2580
aaaaggagct gtatctttta aattggttgc attcaaatcg cttatgctaa tcctattgtt 2640
tttttttttt taatttgtta gcccgtatca tgagtaatta aaaaggaagt taatagccgg 2700
atgtagacaa ctcaatatgc atccgcttgg ttatttgaaa acaattaact tatgaagaga 2760
aatgaatgaa attaccattc ctaattgccc agaagcatga tgcttttgtt aagaagggat 2820
tgttaagaat ttaactatct aaacgaacta atgtgtacat ttctgtgctc cgcgataaag 2880
gaaatgtaat cagcaaaatc aagtttcagg cgcgtggaaa atctgtgaat gaatgaacgt 2940
acgagtaata caaaaagata actaatcacc gctttgcgga tgcaatagcc cagttttgtc 3000
gtcggtgaca ccgagtaaaa acccgatccg atcgacggga ttagcgagga agtctgcgta 3060
ctcccaccag cttatcgcat ccgagacatt tgatttatgt aaacggattt tgttataaga 3120
tttctgccga gccaagctgg ccaactggcg aatcgtaatg atcgagccgg tttgcacgcc 3180
atttgaattc acatttgtac gaattttttt tttatcaaaa gttcgagttt ttcaccaatt 3240
tcctcatcaa ccgagcaagg caaacggctt tgaataatat ggtgttatat atacatatat 3300
caaatcgctg ctgactgcgt gattgatggc cccaagatta catattatcg aatcaggatt 3360
cagaaggaga tcaatgtcaa atgcggacag gaacatgaga gacgcctgtt atgcgcaatt 3420
aaaaatttgg gtttaattgc tgtggaaact gttgttggtg gcatcttaag ttcctgttta 3480
acaacatcaa ctacttatgt acgtagaagc gtttaagcca tttgcataca gatgagaact 3540
ggcttttgtg ctaatcagcc aagatgactc cgatgatgat gactcattac ctgaccagtt 3600
ttcgctgctt tcttttcaac aactacttgt atatgtattg tatccaatag caatacattg 3660
gatttccatg gtctagtcac gtattatcat ttaattgacg ccaagtcgtg ttattgttga 3720
gctatcgagt tcagctcaaa catttcttat tcccatgaat aagccggcaa aaatatgcaa 3780
tctatgaaag ttaatataag caaaccttac tttgactcaa taccaatgca ctttgtgtcg 3840
ataggttcac gcaattgagg cgattattcc gataacccaa gcgattgact gttcccgttt 3900
cgattccaat tgaaatttgg aaatgtacaa tagttttgct atatgctgtc aagtacgctc 3960
ttatcttctc tgggttttct tcagagtttc gaaacgcttc ttcttttttt tgtttttttt 4020
tttttggaat ctcgtatttt ggaaggggct cccctctgga atttgttaca ctgtcgttat 4080
cattgcgaac aagcggcccg aagctatcag cgactttaac atttacaatg cactttttta 4140
cgaccaatta aatgtacatt ttcctttctt cgcccgttga taagcgaacg cgatgtggcg 4200
caggcaatgt gttgctcttg cgacacaaac gcaatcaaaa tggattcaat ttcgcttttt 4260
cccagtgaaa cgaagaacga accgaccatc atgatatgct gctctgcatg ttgcgtattg 4320
aatcaatgac aatttcaatt aagccgcccg ttcgtcatgc gttttcgtgg gcttcgaaat 4380
gctgataacg ctgctgtcct ccaactgctt tgcatgtgga cacaattcca tttatttaat 4440
tcttttattt ggatcggtta aattaaaaag cgccttgtta cgcatttaac gttgtttccg 4500
gtgcgtggtg gtttcatgct tctgggaacg gcaaatgggt ttaggattgg gaacccctca 4560
tcatccgtcc aaatatgggg agttcaagga atggaggggg gagggggga gggggcggag 4620
cagcggtccg ctttcatgca tgatcaacat gtgtgtttgt tttccaactc tatgtgagtt 4680
gtttcagttc ttttcggtct cctcggttct actcagttcg ttcttacttt ctggcttttg 4740
tttccgaata gggggagcgc atcgggccgc aactggaaat ggaaatctga atatgcatcc 4800
cggagcatcc cgtccgaaat gagcaccatg cgccatgccg tttcgtttaa ttccctaccc 4860
ataaatgcgt gtccaaacag cagagggcat ttcaagaggt ccaactatta gctatgtgca 4920
tatataattc tctagatata ttaattcagc taaagtgcta accgcgagtg tgccgccagt 4980
tgcccctcaa cagtttgcaa cgatgatgat ggcacagata tcagagataa cactcagatc 5040
gcgcccctca caagtatgct cattatgtac attcatatgt acatatgtac acacagctaa 5100
gcacataagg aatacgatat gcgctatgta catatataca tatttacatc aaagtcttat 5160
cagcctcggg ttcccaccgt ccaaaaaaaa aggcggaacg aagctttata aataatcaat 5220
caacaaatat ttacaatcaa gtatttatgt agcgggcgga aagtcgaatg caaataaggt 5280
tagacgcatt tgtacatatg tattttctta aagggttcct ctgatcaata tctgcccgtt 5340

tttatctagc cccgtttata aggctgtgaa tgccccgatt gtgttggtcg tgatatcaat 5400
taaatcggca taatctcatg tcatcttttt aaacattatg atagataaaa tgtattataa 5460
cgactctaat tacattggga ttaaatgggt tttcggaaaa tacctatgca gtaaatcaaa 5520
tttaagatca atgtgcagtt atttggggtt tcttgagtta ggacctctgt tttaatattg 5580
agaattactc gatattatct tccgttttat tttattttct gatcgtgtcg gtgcacgaga 5640
aatacatatt tcccactttt gttgcatttc gctgactaat cgaaagctaa ataccaaata 5700
agtttccgat atggggggta tgtatgtaca tacatacata tatagtacga gtatatataa 5760
tatacactga ataagcccga gcatacatat gtacatatgt agatagtgca catgtgtatt 5820
gtggatcgtt ttggagctgg gtttttcttt gcttttcgtc gtttttaaga gacttgagct 5880
gaatgagtcg ctaccaactg ttgcggtggg ttaaagggtg gggggagggg gcgaatgggg 5940
atgggggatg aagatgatga ttcatgcagc tgatgtttgg acagggaaga aagaaagagc 6000
gaaagagaaa gcaaatcgga agcggcggta tcaaacaatc tgtttaaata acaacaaaag 6060
actttataag cctgactttc atttccattt ccaatttatt tttcccaagt ctggctcatt 6120
ttgtgcaaat aaacaaatga ctcatttgtt taaaaattcg atgagactca aggaatgtca 6180
cattcacaaa aattttcaat ttgaaatctg aaaagattca acagcataag caaatgaatt 6240
taagattgcg atatttgggt attatcctta attggctggc tgcctaatgt agatttgttt 6300
ttctcggtgt accgatgcat cctctttgcc catccacttc tttctcccct ctttgctgat 6360
tttttttttt tttgttttgt gagagggtgg aattcgggat ttgggatgtg agcgagaagg 6420
atggccatat tttgaattgt ggaaacagcc gttccattat cgttccatta caaatgcaca 6480
cacacacaca cacacacgcg cgcgtgcagt ttgcgcttct ctgtgtgtgc gtttggacca 6540
aatgcaataa caataaaaac gcgcaactca aattgttaac aaaggtaagg ccgttttttt 6600
ggccggattt ggattctggt cttgaaagag aaacacacac acacacacac gcacgcaaaa 6660
aaccaacgaa aataataaca aaacaacaac tcacccgaac tcgaacccgt ttcactgttt 6720
tccgcttatt ataaatttat acttcgtttg gcgtttctta aatgctttcg aattgcattt 6780
tggatttgac attaaatcgc tccgcgcagc acagtaacaa aaataaacgg aacaaaataa 6840
atacaaccag ccgattctgg ccttctcaca attccgactt ttgaattatc gaaacagaag 6900
cgctgcttgg gccaactaag ccaacttaac taaccgatcg cagagcgcaa agtgtcgtat 6960
tttcactgtc gtcaattaac aacagcggca acaagaacaa caaacgaaat gagcggctgg 7020
gctacagttt tttagtccga gcgcgatcga acatgttggc tgcctggcaa caagttcgat 7080
gcgaactttt ctattgatct aagaaatcgg cgttctatcg attgacgcgc tgcgcaggcg 7140
cctatcgatt tttgaaattc taagcccgcc ttaaaggata tttgtagaaa agtagcttgc 7200
aaagaaatta aattccttga aataaaaaag aaatttagta ctctcgttct atactcactt 7260
ctattcactt gcggttagtt agttagtttt taatgccaca tttttatgat gatgactttt 7320
tttcagctca tctatgcgca atcaatttg cgccaagcga cagctgcaca tagtttcttt 7380
tgaccaagca atcatttatc atataatttt ccaagtagta gtgccaagtt cataactgat 7440
ggagtgtgct gccaatttgc atgttgtgga gccattcagt ttcaaggtcg cgctaattaa 7500
atactcgtta caagtgatgc caaatgtata aatatttgaa cagaacgtat taaattaata 7560
ttatccacta agagcactat tttaaattct taaaaaccag atgcacgtca aatagatgta 7620
cataaaaagt atatttactt acaaatatca tacacggacc tgttttattc ctcatcactg 7680
gccattagtt gataacatgt ggacttctat gatatcctgc tgcatttttg ccgtttagtc 7740
aggatttaga gcttgtacac aatgcgacag ctgaaagggc gaaatcggtg caatcgtgct 7800
gttcggcatc tgaaagttca gggcaagatg tggttgaaaa atcttaaaag cggtttggaa 7860
caaattcgag gtgaatctat ccaaatcgtt agaatcccat gagtacgttc atc 7913

<210> 17
<211> 1399
<212> DNA
<213> Drosophila

<400> 17
ggcccaagca gcgcttctgt ttcgataatt caaaagtcgg aattgtgaga aggccagaat 60
cggctggttg tatttatttt gttccgttta tttttgttac tgtgctgcgc ggagcgattt 120
aatgtcaaat ccaaaatgca attcgaaagc atttaagaaa cgccaaacga agtataaatt 180
tataataagc ggaaaacagt gaaacgggtt cgagttcgga atggatcagg acatgggcat 240
ggcaacgggc tactttcagg acgcagacat gcagatggat gaaccggcgg ccgccacgca 300
atcgatctac ccgcattcgg cgacattatt cgcggccatt agtgcctgtg tctttgtgac 360
gatcggcgtt cttggcaacc tgattaccct gctggcgctg ctcaagagcc ccacgatacg 420
ggagcatgcc acaaccgcct tcgtcatttc gctaagcatc tccgacctgc tcttctgctc 480

```
cttcagcctg ccactgaccg cagtgcgatt cttccaggag agctggactt ttggtaccac 540
attgtgcaag atctttccgg tgatcttcta tggcaatgtg gctgtttcac tcctcagcat 600
ggtgggcatc accctgaaca gatatatact catcgcttgc cacagccgct actcgcagat 660
atataagcct aagttcataa ccctgcagct gttgttcgtt tgggccgtct cctttctgct 720
gttgctgccg cccatattgg gcatctgggg cgagatgggc ctggacgagg ccaccttctc 780
ctgcacaata ctcaagaagg aggggcgatc gatcaagaag acgctgttcg tgatcggctt 840
cctgctgcct tgcctggtga tcatcgtttc gtactcgtgc atctacatta cggtgctgca 900
ccagaaaaag aagattcgca accacgataa cttccagatc gcagcggcca agggctcctc 960
gtcgtccggc ggcggatcct acatgaccac cacatgcacg cgcaaggcgc gcgaagacaa 1020
tcgactgacc gtcatgatgg tcaccatctt tctgtgcttc ctcgtctgct tcctgccgct 1080
aatgctggcc aatgtggtgg acgacgagcg caatacctcg tacccctggc tgcacatcat 1140
cgcctccgtg atggcctggg cttccagtgt cattaacccg atcatctatg cggccagcaa 1200
tcgcaactac agagtggcct actacaagat ctttgcgctg ctcaagttct ggggcgaacc 1260
gttgtcgccg atgccgagta gaaactatca ccaaagcaag aactccaagg agctgtcggg 1320
cgtcatccgc agcacgccgc tgtttcacgc tgtgcagaag aatagtatta accaaatgtg 1380
ccaaacatat tcagtatga                                              1399
```

<210> 18
<211> 386
<212> PRT
<213> Drosophila

<400> 18

```
Met Ala Thr Gly Tyr Phe Gln Asp Ala Asp Met Gln Met Asp Glu Pro
1               5                   10                  15
Ala Ala Ala Thr Gln Ser Ile Tyr Pro His Ser Ala Thr Leu Phe Ala
            20                  25                  30
Ala Ile Ser Ala Cys Val Phe Val Thr Ile Gly Val Leu Gly Asn Leu
        35                  40                  45
Ile Thr Leu Leu Ala Leu Leu Lys Ser Pro Thr Ile Arg Glu His Ala
    50                  55                  60
Thr Thr Ala Phe Val Ile Ser Leu Ser Ile Ser Asp Leu Leu Phe Cys
65                  70                  75                  80
Ser Phe Ser Leu Pro Leu Thr Ala Val Arg Phe Phe Gln Glu Ser Trp
                85                  90                  95
Thr Phe Gly Thr Thr Leu Cys Lys Ile Phe Pro Val Ile Phe Tyr Gly
            100                 105                 110
Asn Val Ala Val Ser Leu Leu Ser Met Val Gly Ile Thr Leu Asn Arg
        115                 120                 125
Tyr Ile Leu Ile Ala Cys His Ser Arg Tyr Ser Gln Ile Tyr Lys Pro
    130                 135                 140
Lys Phe Ile Thr Leu Gln Leu Leu Phe Val Trp Ala Val Ser Phe Leu
145                 150                 155                 160
Leu Leu Leu Pro Pro Ile Leu Gly Ile Trp Gly Glu Met Gly Leu Asp
                165                 170                 175
Glu Ala Thr Phe Ser Cys Thr Ile Leu Lys Lys Glu Gly Arg Ser Ile
            180                 185                 190
Lys Lys Thr Leu Phe Val Ile Gly Phe Leu Leu Pro Cys Leu Val Ile
        195                 200                 205
Ile Val Ser Tyr Ser Cys Ile Tyr Ile Thr Val Leu His Gln Lys Lys
    210                 215                 220
Lys Ile Arg Asn His Asp Asn Phe Gln Ile Ala Ala Ala Lys Gly Ser
225                 230                 235                 240
Ser Ser Ser Gly Gly Gly Ser Tyr Met Thr Thr Thr Cys Thr Arg Lys
                245                 250                 255
Ala Arg Glu Asp Asn Arg Leu Thr Val Met Met Val Thr Ile Phe Leu
            260                 265                 270
Cys Phe Leu Val Cys Phe Leu Pro Leu Met Leu Ala Asn Val Val Asp
        275                 280                 285
Asp Glu Arg Asn Thr Ser Tyr Pro Trp Leu His Ile Ile Ala Ser Val
```

```
    290                 295                 300
Met Ala Trp Ala Ser Ser Val Ile Asn Pro Ile Ile Tyr Ala Ala Ser
305                 310                 315                 320
Asn Arg Asn Tyr Arg Val Ala Tyr Tyr Lys Ile Phe Ala Leu Leu Lys
                325                 330                 335
Phe Trp Gly Glu Pro Leu Ser Pro Met Pro Ser Arg Asn Tyr His Gln
            340                 345                 350
Ser Lys Asn Ser Lys Glu Leu Ser Gly Val Ile Arg Ser Thr Pro Leu
        355                 360                 365
Phe His Ala Val Gln Lys Asn Ser Ile Asn Gln Met Cys Gln Thr Tyr
    370                 375                 380
Ser Val
385
```

<210> 19
<211> 3574
<212> DNA
<213> Drosophila

<400> 19

```
gttcgtttga tcggttgacc tcgatttatt tactgctatg caaagctaaa attcaaaaat 60
ccctactcat agatgaagtt ttttgcacac tttaggtgta acttggtgct ttaacttggc 120
gtttctagtc ttaaataatt ggcaagatta tctgcctaac ttttggtgac cccaactttt 180
aatcacacac caagcggaat tatactaagg tttatttccc ttggcatcag actgtcttct 240
tcttcgctct tccagctctt gtctttatta aacctggcca aggacctgta acttcgttta 300
attaagtaat tcaatgaaat tgacagctgt ttgcttgggt caagcataat tgatgactac 360
tcataatcaa aacagaagtg ctgccaatga atgggaagcg gaagtctggc ataaattaca 420
ggtctgtcca tccgaagcgc aatcatcata taattgtcat ggaaccacac ttggcactca 480
cttttgcctg gcttctgctc ctgcgaacgc tgatgttgct ttgtgctttg tgctttctgc 540
ttgatgcttt ttgctttgtg cccaagtctc gtgactccac tccactcgga gtggcatcca 600
tggttaggag aacagcaata atgaggcaca ccaaccggcc aattggcaat tgtgcgaggt 660
caaacacaac cggtggagga aactggtgtc ggccgcacag ggaagactta aatttctttt 720
aattgctcga atttatgggt tggtagtggc ataacgacaa ttgtccaagt caaggtcaca 780
aatgatgctc ctgccttttt agttcgatac acacaaagaa aaggggacca tacatataca 840
agcaaccaac tagatgaaac ctattttaag gacggttgaa ttgttggctt ttaacatagt 900
agattaactt caaaattagc agttactact actaataacg agctaagcca ctcatttgcg 960
taaacacttt ctatctgtgc acttttatct gcttgaatag tgtctgggtg atggcctcga 1020
ttgccgtggt tggcaaccta ttggtcctgc tcggtcgcta cttctacaaa tcacggagca 1080
acgtggagca ctcgctctac cttcgccatt tggccgccag tgatttcctc atgggcattt 1140
acctcacatt gattgcctgt gcggatatca gttttcgcgg cgagtatatc aaatacgagg 1200
agacctggcg gcatagcggc gtttgtgcct tcgcaggtgc gtatccttcg acatcctagc 1260
catttccatt tcgtaatgga atgctaatgt gtattacgta ttcgccgtgt aggcttcctc 1320
agcaccttca gctgccagtc gtcgacgctg ctgctcacat tggtcacctg ggatcgtctg 1380
atgtcggtga cgaggcccct caagccacgg gatacggaaa aagttcggta agtgtgaact 1440
acaatagata tgtaataaag tgctgggcaa agttattgtt tattctatta cattaattac 1500
ttttgctgcg gacatttgct gatttgttga ctgctcaaat atcaatgcac agaatttgtt 1560
taccctactt gaattttgtt cgtaaatgca attatgtttg ttttgaccgc cgccttacta 1620
tgtacaatat atattttctg ttgcttcatt ttagtttgca aaagctatgg ttttgtatgt 1680
gtaagctatg ataaattatt gaaatgacaa cttccacagc attgtcctac gccttctgct 1740
gttgtggggc ataagtttcg gattagctgc agctccactt cttcccaatc cgtactttgg 1800
tagccatttc tacggcaaca atggtgtctg cttatcgctg catatccacg atccctatgc 1860
gaaggtaata caccatttcc ttggaaacgc ctgttctaat attccaaaat gtatattttt 1920
agggttggga gtactcggcg ctgctgttca tcctggtcaa tacgttgtca ctgatcttca 1980
tcctcttttc ctacattcga atgttgcaag cgataaggga ttcgggtggc ggaatgcgaa 2040
gcactcacag cggtcgcgag aatgtggtag caactcggta agtttggata attcaagtaa 2100
aatgatttaa ttattgaaat atcatgtttt tacagctttg ccatcattgt gaccaccgat 2160
tgcgcctgct ggctgcccat aattgtggtc aaattggctg ctctttcagg caagttcatc 2220
ggcaacgcct ttaaatgatt tattttgatt atcaacttat tcgtatatat ataacccctt 2280
tcaggctgcg aaatctcgcc cgatctttac gcctggctgg cggttcttgt gctgccagtg 2340
```

aactcggccc tcaatccagt gctctacaca ttgaccactg cggcctttaa gcaacagctg 2400
cgtcgctact gtcacaccct gcccagctgc tcgctggtga acaacgagac ccgatcccag 2460
acccagactg cctacgagtc cggactgagt gtcagtctgg cgcacttggg cggtggtgtg 2520
ggcggcgggt cggggcgaaa gcggatgtca caccggcaga tgagctatct gtaggcgggc 2580
gcaactggct taaaatgcag cacgagatca acagactggg gctagtgggt gggtggcatt 2640
acaacaggac acttccatgt aattggtccc ttagactatt ggtattatac gtacttattg 2700
atacttattc taatatcaat tttaaattac ttataacatt tgaggatatg ataacataaa 2760
ccttagacaa gagattccag ctatttaagt ttcttttttg tatattattt attttaccaa 2820
tcgttttaag gcaaacctct tacaactgtc caaagaatat cctaactatt tgttgcatcc 2880
ccatttattc gacctgccca atgggatatt gttacatgtg atctaatagt taaatcaaat 2940
tttaatttta taaatacata ttcagatgtt aattgaagct tttgaacata ggacttgcaa 3000
tcgaaggccg caaatgcctg ccatccggct ttctatgcgt tgccctgagg ttcagtggct 3060
acatcgagat ccgcctggac gtgatagggt ccctcgttca cataggccgt gaaacggatg 3120
gtttcggtgg gcgaaaactc agcgctagcg gacagaatct ccataatttg ccacttcata 3180
tacttaatgt ccctgtgcag atccttttga cgcgaaagaa ggaactccac agcaccgccg 3240
gtattcttta gagctacaat gccccagttg tacttacaaa tggttttcaa gaacgacaag 3300
gcggccattt gaagatcggg gaacggagtg tttatcaggt ccataatgat attagcatgc 3360
gctcccccag cgaagcactc gtaccagttt ttaccaatgt tactgaaaaa gaagagaatc 3420
attagaagaa aaacttcaat tatagagtta agcattacat gatttctgtg gcaggcggcg 3480
tcttgaagtc ataaatcaca tccagggagt tgagcagcct cttcttgatg tgcgccgata 3540
gcttcttagt atgcgagcca tacttcttaa agga 3574

<210> 20
<211> 1015
<212> DNA
<213> Drosophila

<400> 20
tgtctgggtg atggcctcga ttgccgtggt tggcaaccta ttggtcctgc tcggtcgcta 60
cttctacaaa tcacggagca acgtggagca ctcgctctac cttcgccatt tggccgccag 120
tgatttcctc atgggcattt acctcacatt gattgcctgt gcggatatca gttttcgcgg 180
cgagtatatc aaatacgagg agacctggcg gcatagcggc gtttgtgcct tcgcaggctt 240
cctcagcacc ttcagctgcc agtcgtcgac gctgctgctc acattggtca cctgggatcg 300
tctgatgtcg gtgacgaggc ccctcaagcc acgggatacg gaaaaagttc gcattgtcct 360
acgccttctg ctgttgtggg gcataagttt cggattagct gcagctccac ttcttcccaa 420
tccgtacttt ggtagccatt tctacggcaa caatggtgtc tgcttatcgc tgcatatcca 480
cgatccctat gcgaagggtt gggagtactc ggcgctgctg ttcatcctgg tcaatacgtt 540
gtcactgatc ttcatcctct tttcctacat tcgaatgttg caagcgataa gggattcggg 600
tggcggaatg cgaagcactc acagcggtcg cgagaatgtg gtagcaactc gctttgccat 660
cattgtgacc accgattgcg cctgctggct gcccataatt gtggtcaaat tggctgctct 720
ttcaggctgc gaaatctcgc ccgatcttta cgcctggctg gcggttcttg tgctgccagt 780
gaactcggcc ctcaatccag tgctctacac attgaccact gcggccttta agcaacagct 840
gcgtcgctac tgtcacaccc tgcccagctg ctcgctggtg aacaacgaga cccgatccca 900
gacccagact gcctacgagt ccggactgag tgtcagtctg gcgcacttgg gcggtggtgt 960
gggcggcggg tcggggcgaa agcggatgtc acaccggcag atgagctatc tgtag 1015

<210> 21
<211> 334
<212> PRT
<213> Drosophila

<400> 21
Met Ala Ser Ile Ala Val Val Gly Asn Leu Leu Val Leu Leu Gly Arg
1 5 10 15
Tyr Phe Tyr Lys Ser Arg Ser Asn Val Glu His Ser Leu Tyr Leu Arg
20 25 30
His Leu Ala Ala Ser Asp Phe Leu Met Gly Ile Tyr Leu Thr Leu Ile
35 40 45
Ala Cys Ala Asp Ile Ser Phe Arg Gly Glu Tyr Ile Lys Tyr Glu Glu
50 55 60

```
Thr Trp Arg His Ser Gly Val Cys Ala Phe Ala Gly Phe Leu Ser Thr
65                  70                  75                  80
Phe Ser Cys Gln Ser Ser Thr Leu Leu Leu Thr Leu Val Thr Trp Asp
                85                  90                  95
Arg Leu Met Ser Val Thr Arg Pro Leu Lys Pro Arg Asp Thr Glu Lys
            100                 105                 110
Val Arg Ile Val Leu Arg Leu Leu Leu Leu Trp Gly Ile Ser Phe Gly
        115                 120                 125
Leu Ala Ala Ala Pro Leu Leu Pro Asn Pro Tyr Phe Gly Ser His Phe
    130                 135                 140
Tyr Gly Asn Asn Gly Val Cys Leu Ser Leu His Ile His Asp Pro Tyr
145                 150                 155                 160
Ala Lys Gly Trp Glu Tyr Ser Ala Leu Leu Phe Ile Leu Val Asn Thr
                165                 170                 175
Leu Ser Leu Ile Phe Ile Leu Phe Ser Tyr Ile Arg Met Leu Gln Ala
            180                 185                 190
Ile Arg Asp Ser Gly Gly Gly Met Arg Ser Thr His Ser Gly Arg Glu
        195                 200                 205
Asn Val Val Ala Thr Arg Phe Ala Ile Ile Val Thr Thr Asp Cys Ala
    210                 215                 220
Cys Trp Leu Pro Ile Ile Val Val Lys Leu Ala Ala Leu Ser Gly Cys
225                 230                 235                 240
Glu Ile Ser Pro Asp Leu Tyr Ala Trp Leu Ala Val Leu Val Leu Pro
                245                 250                 255
Val Asn Ser Ala Leu Asn Pro Val Leu Tyr Thr Leu Thr Thr Ala Ala
            260                 265                 270
Phe Lys Gln Gln Leu Arg Arg Tyr Cys His Thr Leu Pro Ser Cys Ser
        275                 280                 285
Leu Val Asn Asn Glu Thr Arg Ser Gln Thr Gln Thr Ala Tyr Glu Ser
    290                 295                 300
Gly Leu Ser Val Ser Leu Ala His Leu Gly Gly Gly Val Gly Gly Gly
305                 310                 315                 320
Ser Gly Arg Lys Arg Met Ser His Arg Gln Met Ser Tyr Leu
                325                 330
```

<210> 22
<211> 8412
<212> DNA
<213> Drosophila

<400> 22

```
ttttccgcca aattcgctgg cttttaagtg ctcacataaa tgaaattttc aaatagcaaa 60
taacatagga ataatccttc taaatcgcaa tataaaaaca tatatacaag gagtggtaaa 120
tgacagtatt ctattaatgg caaaacgatt tatttaaaca gacgtcgcat ccgcataaac 180
gatctttcac actttttatt aacggttagt ataaataaaa aaaaacatta atccttagtt 240
tatttaacca attttgtttg ctgaaatcga ttcttgaaag agaagctgac attttttttta 300
tatatgcctc cctacttcag gcagtgcgaa cataccgcat caacgcgact tgtagcggcc 360
gtgtgaaaag gctattttgt ggttagcgca gatctgcgct gcagtgtgac cccggctaca 420
ggtatttcgg catcgcacag ggtcacactc gcccaggcac ccgtataaaa tccccgcccc 480
cctgatcgac gatagtattt cgaaaaaaac gcgaggaaaa gtctgatgcg cgttaaaatc 540
ggagcgcaaa tcgtgcatat attttttttt acacaccgcg cgccgtgttt acgactgaaa 600
acggagactg actgaaaagc cttggcgcac aagcacgcaa tgaaaacgca acgataagaa 660
tgtgatattg tgataaatac atgattcgct gacaactcta ttttatattg gaatgcattc 720
gaaatcgaac gcagcgcagc gtgtgttgtt tacttaacga ttttcggtgt ttgtttgccc 780
aacttgcggc ggtgaaattc agcattttct gattttccat ttgccaacga aagagaggaa 840
aagtgaaaag tgctattcga cgaagcgaaa cggaacgcaa aatcaaaaga aaaacagcat 900
agctaaagct atgtgtagta catagccaca taaattgcaa attgaggcca agagacaaga 960
gaacctgaaa tccacacaga tcttgacgag tgagctaacg atggacagga gtcggagtag 1020
cagagctgcc agcagcaacc agttcatcag gccatgtggc ctactgacga ccgtcatcct 1080
```

gctgcagacg cttgtgtcga tgagtttggc catcgaggag atgtcgccac caccggccgc 1140
cccgccacgt ccatcgccgc cgcccaccgt gaagctcaac aagtgctgcc attcgggcga 1200
gtatctgaac gatggcacct gcatcgccgg cagcgaagcc ctgtggctgc ccatggtgta 1260
cttggtgcag cagcagcgct tcttcgagcc ccatggcgcc agtccgcgat tcttgaagtt 1320
cctgcccaat acgagaccca cgtgccgaaa ggaccagacg acggagatct ttcgcagccg 1380
cggagccaat gtgatgctct ttcccaatgg cacgctctat gtgcgggaac gcgctctaat 1440
ggtccagccg tcggattact gtgtcgactg ggaggtggcc gtcgtttgct tgaacgacag 1500
tcagccgatc aatgctctgg aggatcccga ttatgcagcc aatcctctcg tgcagcagga 1560
gccgccсaag gctagcctgc gactgagcaa gtgctgcggc aagtggggca gctataacac 1620
gcagctccag aactgcgatc tgcagcccaa ccatcaggca gccgtcgatg gactgctgcg 1680
attgtcgccg cagctaccgg agggcagcta tcagactagc tacggcctgc ccgactgtgg 1740
ccaacccggt ggctactcca ttgccggtga ttggcaggat gccaagctgg atcggaacac 1800
ggccatgctc cagttgccgc acaagaacct cagcgccgga caatactgcc tggagcacac 1860
gcagcgcgag ggcgaggtga agatcatagc ctgccagcat ttgtttagct ccgcggcagg 1920
agccggcatc cacgacggct ccattggagg caccatcgag caggccaatg gccagaatct 1980
gcagaaggcc gtcctgactg gtggcatact tgtgtccatt gtcttcctgt ccgccacttt 2040
ggtggccggc ttcctgctgc ccgctgtgca tcatgcgctt cattggcgct gccaaatctg 2100
ctatgtcacc tgcttgctat tcggcaaaat actgctggcc atcgaggagc tgagctccag 2160
cctgcagccg ggcagtgccg cctgccatac gcttgctatc accatgcagt tcttctttct 2220
ggccgccttc ttttggctga acaccatgtg cttcaacatc tggtggacgt tccgggactt 2280
tcggcccagt tcgttggaac gcaatcagga ggcactgcgt cgctatctgt actccctgta 2340
cgcttggggc ggtcctctgc ttatcacctt cgtggccgcc tgcgtggacc agctgccgga 2400
gacgacgcta ctgcgtccgg gattcggaca gctctactgc tggttcgaca atcgcaatct 2460
ctcgatcttc gcctacttct acggacccat tggcctgctg ctgtgcgcca atatagcgct 2520
cttcgtgtcc accacccatc agttgacgtg cggcctgtgg aaacgggacg atgtcaagtc 2580
ctcatcggag aagtccgccc tgggacgcgt ctgcctgaag ctagtcgtcg ttatgggtgt 2640
cacctggata gcggacattc tgtcctggct agtgggtgga ccgcatggcg tatggttctt 2700
caccgacctg atcaacgccc tgcaaggcgt cttcatcttc atcgtggtgg gctgccagcc 2760
gcaggtgtgg accgcctgcc gcaggatctt ctgtccgcga ttgcgtcatg acatcaccaa 2820
tacgaccaac ggtgtccagc attcgagcag ctcgcagggt ttgccctcga tggccggcgg 2880
cacggagatt acacagaaca ccaccaccac caccacgaca accaacacga ccgccactca 2940
catgcccagc aatccggcgg aggatgaggt accggagaag gctcccattg cccccgtcgc 3000
acccatcgtc aaaatggaga ccatttgcta aacgtaagtt gaacacccga tagccaatga 3060
agtgtgcttg taaagagata gctctgtttg ccagcgcgca aacaaagcgc ctgctgttcc 3120
cactcccagc gcagtttgcc aagtggctta caaaacgcgg ctttaagcac acacacacac 3180
acacacacac ttgctgatag acttatcggc cttaatgggc tttagggccc agccgggaat 3240
gcaaagtatc tcaatgagct ggctggctgg ctggctggct gtttgctcgc cggttacgcc 3300
atgcatcgca gcagtttcga tcctataaag cgcacgtatg tacaggtggg gccacttcaa 3360
tatcactaag cagctttagc ttcaatagac ttcaatcggt ggaaggtttg gtatttggac 3420
ccaagaacct aagtttctta ttaagtttta gtctttcaaa aatagcgtgc taaatagagg 3480
gtcttaaatt aactcgaaaa tgatttagaa tccgtgtaaa ttgaattgac acattagctg 3540
aattttccgg tatataacca aaaatgtccg gacgatatat gtgtcgtaga agatagtgtg 3600
ataattgcgc tgtctgtgga tttaacccaa ataccaatga ggtattcaca aaatgtgaaa 3660
acaaatagga aaccccgaat ttgtcatgcg ccgacaactt tcgagaacca accaccacca 3720
tatgaagata aaaaacaaaa gaaactagtt tggttttcta tggggaaaa cgatacgaac 3780
ttttcaattc aaaacctttg gagtttcaac ggtcaatggc catgttggaa aatacccttc 3840
tctataggac actattcatc tcacaccccc catgtggact atacgctatc cagatgaaat 3900
ttaagggttc gagatttcct gacaacttgc ggctacctgg agcgtttttc caaccaggcg 3960
ccggttgaaa aagtctagca cgttttcgtt tcgttttatg gttttatagg gggaaattta 4020
cgtatttacg taggataaga agatcgtcct gcgatggtca aatagtgacc atattcttca 4080
tttgcatgac cgaaattgtt taacaacatg cgccaagttt acgatttcgc gtgcaccact 4140
tgagttcgag tatggtcata aagatgaaaa tgttgccgct gtgtgccgtt atgtttttca 4200
tcggaaaact ctgacgccgg gccgggcgaa ttcgattttt gaatttcctg tcgatttcca 4260
catttccaca tggccataaa gtaatcagcg caacccttga cacacacgca agcgaagata 4320
atggaaatcg ttgctggcat ccgtatcggt tggattcaaa aaatgtgtgt tgccgaatcc 4380
gtaagggaaa atcgagttta aattggactc accggcgacg atggcgatgg tgatggtgat 4440
aagcacacaa cacaccatcg tcttgtttgg atttcggccc atttacttat ttacttgtta 4500
ctttgttagt ttgttagttt gttagtttgg tcgcctcgag tggttttgac acatctccca 4560
atgatccaa tggcaaacat gcggacaggt ctcgaactcg gcacaatcta caattatcta 4620
aagaaacc tctgtgcggc gatcataaag ttttgaaaaa caataattcc attttcggc 4680

actaaaaaaa aaacagggca acttgaaatc tagcattaaa ctggtggctt aatgttccaa 4740
tgaaaagata aggatttaac aataaaaaaa aacttttttt ttgtatgcac tgcatgttga 4800
tagcaacatt aatgaaagaa actaatcaat gaagtttctg tttctctata ttaacaacaa 4860
atgctgcttt gtttaggcca tgtttttgtt attagagccc aagaaacatt ggtttgcatt 4920
gcacaaaggc gtatggaaaa tgtgtaatta cttcagtttg taacattttc gaatatgaca 4980
gctgctggtg gcagttaatt ctttgccctc gcgttttatt gtttgttact ttttactttt 5040
atttttttatt ttttttttg ggccctgatc gcgtcatact ttttgggatg cgtgtctgtc 5100
cgggtgaaca gatgtgctgt ctgtctgcca ttcgcttgtg tgtaggtccg tgttggaatc 5160
agaggttttg cccaaacgaa agagaaaaca aaacaaacag tgggtttttg gcccatgggc 5220
ccattcgatg agattgcgca accgagttaa atggcctggc ccctgaaaag ctgaaaaccg 5280
aaaacagaaa catgaaatct gaaatctgaa attgtggaaa tctgaaattg tggctggcac 5340
acgcaacact cgtaatcaga ctaaaaaaaa aaggggcagt cagtcggctt aaaactgtga 5400
cctgcatgca tttgccaact gcagactgca aactgcaaac tgcaacttcg aaggcaaccc 5460
gaaatgcatc tgcaactgca actgcagctg ccgtgccatt tgtgttcatc tacataatcc 5520
cagagcagcg tgtccacgtc acgctcatca caacgcgacg gtgcaaaaat gcagcgaaat 5580
tggatgcgtt gctcgtttgt caactgaaat ggtccagatg cacaatgcaa gggagggaag 5640
atgcggcaga aatgcgccgg aatcgatgca gtaatgcgtt tcaagctatc attatgggat 5700
cacatggggt cacatcagtg cgaaaatgct aatgctatat gggctaaaac cagttgggaa 5760
ttagagcgcc gtacaagctg taagctaagc tgatatgcaa taccgcttat taaagagcat 5820
ttaaactatt tgacttgaaa atatgactaa caaattgagc agggaaaatg cttttccttc 5880
ggaaaaaatg attgtgtgcc aaaaactaaa aacgaaagtt tccctactat tacttttggc 5940
atcgcccaaa aaacactttc tttcattagc catttcactg attatctttt ttggtaacca 6000
agcggagtaa tcatcataag cacacagcgt taaactattc agctatttcc agcttgttga 6060
gataactaaa agctcggaca cgcttgagtt ccactttttt tttttttttt ttatctccat 6120
caacatttaa taattaagtg atttcgctta ttgttctagt atttattagc gcattaccta 6180
attgtttgac acgacgaaac atgaaatgat tactaatttc atcatagaat gccagtgacc 6240
catgctcgtt ggcatttcga ttataataat attcgcagtt gcataaatat agattgtttc 6300
cgaacgtgac ctgtacttcc tagtgcaact gaatcttttg cgtttctggg tcagctttta 6360
tctgcctagt tcgctgacat ttcaaaagtt tgcaacactt ggcagcttgg cagcgtaatg 6420
aattccttcc tgttgactcc gccacgcaaa ttggcacaaa tttcgcaata agaatactta 6480
atgtccaact ccagacgcaa tcgcagcggc acacgcacac atgctctgcg ttaatttgag 6540
tttagcggca atttttaatt tgaacgttct gcagtttttc aacttataat tgcactctgg 6600
cccatttaca ttttacgcag ggcattatta ttatttcttt tttttttttt aatttttgcc 6660
tgctgagtcg gtcgacatcg gattcctcgt tggaggccga aggaagcgaa gtagggggca 6720
ttgttgtcat gcatccacag ttgcggttaa tgcaatagct ctgagatgga taataattac 6780
ctcggtcctt gaatggtgcg aaaatatctt ggcagttcct ataaaagcat agcattgcgt 6840
ttggtacatg aatttcattg gcaaataata agtatttcat ctggcagcag ttcacacctc 6900
aaatcatatc gcgcaaatgc caccttttcg caagaatttg aaatcaaaat aaatctcaga 6960
attcctgatt acgcgatgcc tggaaatcta aacagatcat gagtaagccg caaatgtttc 7020
agctaagaaa aatgaaaaga acaaagtttc cgctaagggt agttggaaaa gtagtcaaat 7080
tggtttactg atcaatatta cttttagatc tctctgttgt tttatttatt tagcagtgta 7140
aactaagact gcaagtccat gaactatagc cattttcaac actgcttcta agtactaatg 7200
tcttgagtgc actgtaccca catgcgcacg catatcatat gaaaatttgt acttatttgt 7260
gccccttttc tttcttattt tttttttgc agacaactga gacgacgagc gacaatgcga 7320
aagtcaggcc cggatacaaa agtccagacc tcgagcgagc gaaataatgg caatgggcag 7380
ctcagaaata agaaagaaag aatggaaaaa acacgaaaaa aaaacgctgc caaagctgga 7440
gtttcgagac gagcgcggca cacacttgcg caatcgcagc aacaaaacca agctgaccag 7500
ctaattgtgc gagtgtacga gcgaatgagt gtgtgtgtgg ttgtgtgtgt gtgcaagcga 7560
cttgcgcgtg attttgaaat cggccaacaa ccgctgggga acaaaacgta ctcagtatgc 7620
cagtattcaa cttcaattta agccagcgca caaaagaaac gtattatttt attcatcatc 7680
atcatcataa tcgcatagca tcgcatcgca tcgcatcgca tcgtatcgca gaataacggt 7740
aaagcagttc gggtcacatt aacacgttct aattggacac tccggtcagt gttgcgaaat 7800
ggaatgtgca gcataacgaa acaggaattc agcctagaca aaacaaagca gcgctattgt 7860
ggctctttct ttcttttaag cacttttagc caacagaaat aaaaacacga gggctttatg 7920
tgtgtgcatc tggctgtatg tatgtgtgtg gcaaccacaa ggaattttct gttgaggcac 7980
ttttgttgtt tagtttagat cgcgttgctc tggttttttt ttttttactc aaatctcaag 8040
tgtgagaaaa acagcaggcg caccacacac aaacacacac acacatgcat gcacatatcg 8100
gatgtttaac atttaaatta gctgcttctg attggcattt tcggtatgac actgcttgca 8160
ttcagttctt tatgtgtggg tgtgtgtgtg tgtgtgcctt tgacgttttt attgcctaat 8220
cacttttttt tatttgcact ttactacgag aatactggac gaaaaccgct gactggaact 8280

atttttatgt ttgccccata atgtttttca accgaaaaat tgtatatttt ttttttacgt 8340
ctagtgtaaa caattgaaat ttatttactt cagggcataa acaaaatcgt tatacacaca 8400
cattgatagg cg 8412

<210> 23
<211> 2153
<212> DNA
<213> Drosophila

<400> 23
atggacagga gtcggagtag cagagctgcc agcagcaacc agttcatcag gccatgtggc 60
ctactgacga ccgtcatcct gctgcagacg cttgtgtcga tgagtttggc catcgaggag 120
atgtcgccac caccggccgc cccgccacgt ccatcgccgc cgcccaccgt gaagctcaac 180
aagtgctgcc attcgggcga gtatctgaac gatggcacct gcatcgccgg cagcgaagcc 240
ctgtggctgc ccatggtgta cttggtgcag cagcagcgct tcttcgagcc ccatggcgcc 300
agtccgcgat tcttgaagtt cctgcccaat acgagaccca cgtgccgaaa ggaccagacg 360
acggagatct ttcgcagccg cggagccaat gtgatgctct ttcccaatgg cacgctctat 420
gtgcgggaac gcgctctaat ggtccagccg tcggattact gtgtcgactg ggaggtggcc 480
gtcgtttgct tgaacgacag tcagccgatc aatgctctgg aggatcccga ttatgcagcc 540
aatcctctcg tgcagcagga gccgcccaag gctagcctgc gactgagcaa gtgctgcggc 600
aagtggggca gctataacac gcagctccag aactgcgatc tgcagcccaa ccatcaggca 660
gccgtcgatg gactgctgcg attgtcgccg cagctaccgg agggcagcta tcagactagc 720
tacggcctgc ccgactgtgg ccaacccggt ggctactcca ttgccggtga ttggcaggat 780
gccaagctgg atcggaacac ggccatgctc cagttgccgc acaagaacct cagcgccgga 840
caatactgcc tggagcacac gcagcgcgag ggcgaggtga agatcatagc ctgccagcat 900
ttgtttagct ccgcggcagg agccggcatc cacgacggct ccattggagg caccatcgag 960
caggccaatg gccagaatct gcagaaggcc gtcctgactg gtggcatact tgtgtccatt 1020
gtcttcctgt ccgccacttt ggtggccggc ttcctgctgc ccgctgtgca tcatgcgctt 1080
cattggcgct gccaaatctg ctatgtcacc tgcttgctat tcggcaaaat actgctggcc 1140
atcgaggagc tgagctccag cctgcagccg ggcagtgccg cctgccatac gcttgctatc 1200
accatgcagt tcttctttct ggccgccttc ttttggctga acaccatgtg cttcaacatc 1260
tggtggacgt tccgggactt tcggcccagt tcgttggaac gcaatcagga ggcactgcgt 1320
cgctatctgt actccctgta cgcttggggc ggtcctctgc ttatcacctt cgtggccgcc 1380
tgcgtggacc agctgccgga gacgacgcta ctgcgtccgg gattcggaca gctctactgc 1440
tggttcgaca atcgcaatct ctcgatcttc gcctacttct acggacccat tggcctgctg 1500
ctgtgcgcca atatagcgct cttcgtgtcc accacccatc agttgacgtg cggcctgtgg 1560
aaacgggacg atgtcaagtc ctcatcggag aagtccgccc tgggacgcgt ctgcctgaag 1620
ctagtcgtcg ttatgggtgt cacctggata gcggacattc tgtcctggct agtgggtgga 1680
ccgcatggcg tatggttctt caccgacctg atcaacgccc tgcaaggcgt cttcatcttc 1740
atcgtggtgg gctgccagcc gcaggtgtgg accgcctgcc gcaggatctt ctgtccgcga 1800
ttgcgtcatg acatcaccaa tacgaccaac ggtgtccagc attcgagcag ctcgcagggt 1860
ttgccctcga tggccggcgg cacggagatt acacagaaca ccaccaccac caccacgaca 1920
accaacacga ccgccactca catgcccagc aatccggcgg aggatgaggt accggagaag 1980
gctcccattg cccccgtcgc acccatcgtc aaaatggaga ccatttgcta aacacaactg 2040
agacgacgag cgacaatgcg aaagtcaggc ccggatacaa aagtccagac ctcgagcgag 2100
cgaaataatg gcaatgggca gctcagaaat aagaaagaaa gaatggaaaa aac 2153

<210> 24
<211> 676
<212> PRT
<213> Drosophila

<400> 24
Met Asp Arg Ser Arg Ser Ser Arg Ala Ala Ser Ser Asn Gln Phe Ile
1 5 10 15
Arg Pro Cys Gly Leu Leu Thr Thr Val Ile Leu Leu Gln Thr Leu Val
20 25 30
Ser Met Ser Leu Ala Ile Glu Glu Met Ser Pro Pro Pro Ala Ala Pro

```
        35                  40                  45
Pro Arg Pro Ser Pro Pro Pro Thr Val Lys Leu Asn Lys Cys Cys His
    50                  55                  60
Ser Gly Glu Tyr Leu Asn Asp Gly Thr Cys Ile Ala Gly Ser Glu Ala
65                  70                  75                  80
Leu Trp Leu Pro Met Val Tyr Leu Val Gln Gln Gln Arg Phe Phe Glu
                85                  90                  95
Pro His Gly Ala Ser Pro Arg Phe Leu Lys Phe Leu Pro Asn Thr Arg
            100                 105                 110
Pro Thr Cys Arg Lys Asp Gln Thr Thr Glu Ile Phe Arg Ser Arg Gly
        115                 120                 125
Ala Asn Val Met Leu Phe Pro Asn Gly Thr Leu Tyr Val Arg Glu Arg
    130                 135                 140
Ala Leu Met Val Gln Pro Ser Asp Tyr Cys Val Asp Trp Glu Val Ala
145                 150                 155                 160
Val Val Cys Leu Asn Asp Ser Gln Pro Ile Asn Ala Leu Glu Asp Pro
                165                 170                 175
Asp Tyr Ala Ala Asn Pro Leu Val Gln Gln Glu Pro Pro Lys Ala Ser
            180                 185                 190
Leu Arg Leu Ser Lys Cys Cys Gly Lys Trp Gly Ser Tyr Asn Thr Gln
        195                 200                 205
Leu Gln Asn Cys Asp Leu Gln Pro Asn His Gln Ala Ala Val Asp Gly
    210                 215                 220
Leu Leu Arg Leu Ser Pro Gln Leu Pro Glu Gly Ser Tyr Gln Thr Ser
225                 230                 235                 240
Tyr Gly Leu Pro Asp Cys Gly Gln Pro Gly Gly Tyr Ser Ile Ala Gly
                245                 250                 255
Asp Trp Gln Asp Ala Lys Leu Asp Arg Asn Thr Ala Met Leu Gln Leu
            260                 265                 270
Pro His Lys Asn Leu Ser Ala Gly Gln Tyr Cys Leu Glu His Thr Gln
        275                 280                 285
Arg Glu Gly Glu Val Lys Ile Ile Ala Cys Gln His Leu Phe Ser Ser
    290                 295                 300
Ala Ala Gly Ala Gly Ile His Asp Gly Ser Ile Gly Gly Thr Ile Glu
305                 310                 315                 320
Gln Ala Asn Gly Gln Asn Leu Gln Lys Ala Val Leu Thr Gly Gly Ile
                325                 330                 335
Leu Val Ser Ile Val Phe Leu Ser Ala Thr Leu Val Ala Gly Phe Leu
            340                 345                 350
Leu Pro Ala Val His His Ala Leu His Trp Arg Cys Gln Ile Cys Tyr
        355                 360                 365
Val Thr Cys Leu Leu Phe Gly Lys Ile Leu Leu Ala Ile Glu Glu Leu
    370                 375                 380
Ser Ser Ser Leu Gln Pro Gly Ser Ala Ala Cys His Thr Leu Ala Ile
385                 390                 395                 400
Thr Met Gln Phe Phe Phe Leu Ala Ala Phe Phe Trp Leu Asn Thr Met
                405                 410                 415
Cys Phe Asn Ile Trp Trp Thr Phe Arg Asp Phe Arg Pro Ser Ser Leu
            420                 425                 430
Glu Arg Asn Gln Glu Ala Leu Arg Arg Tyr Leu Tyr Ser Leu Tyr Ala
        435                 440                 445
Trp Gly Gly Pro Leu Leu Ile Thr Phe Val Ala Ala Cys Val Asp Gln
    450                 455                 460
Leu Pro Glu Thr Thr Leu Leu Arg Pro Gly Phe Gly Gln Leu Tyr Cys
465                 470                 475                 480
Trp Phe Asp Asn Arg Asn Leu Ser Ile Phe Ala Tyr Phe Tyr Gly Pro
                485                 490                 495
Ile Gly Leu Leu Leu Cys Ala Asn Ile Ala Leu Phe Val Ser Thr Thr
            500                 505                 510
His Gln Leu Thr Cys Gly Leu Trp Lys Arg Asp Asp Val Lys Ser Ser
```

```
        515                 520                 525
Ser Glu Lys Ser Ala Leu Gly Arg Val Cys Leu Lys Leu Val Val Val
    530                 535                 540
Met Gly Val Thr Trp Ile Ala Asp Ile Leu Ser Trp Leu Val Gly Gly
545                 550                 555                 560
Pro His Gly Val Trp Phe Phe Thr Asp Leu Ile Asn Ala Leu Gln Gly
                565                 570                 575
Val Phe Ile Phe Ile Val Val Gly Cys Gln Pro Gln Val Trp Thr Ala
            580                 585                 590
Cys Arg Arg Ile Phe Cys Pro Arg Leu Arg His Asp Ile Thr Asn Thr
        595                 600                 605
Thr Asn Gly Val Gln His Ser Ser Ser Ser Gln Gly Leu Pro Ser Met
    610                 615                 620
Ala Gly Gly Thr Glu Ile Thr Gln Asn Thr Thr Thr Thr Thr Thr Thr
625                 630                 635                 640
Thr Asn Thr Thr Ala Thr His Met Pro Ser Asn Pro Ala Glu Asp Glu
                645                 650                 655
Val Pro Glu Lys Ala Pro Ile Ala Pro Val Ala Pro Ile Val Lys Met
            660                 665                 670
Glu Thr Ile Cys
        675


<210> 25
<211> 3187
<212> DNA
<213> Drosophila

<400> 25
gaccagacaa gtgcaagcaa tggagctcag cgtcagtatc cactgtctca cacacatcca 60
caaaagcatc ttcatcagat tcagttcctc agatctcgca tcagctgctg cgattcccat 120
ggacgtgaca cggtggagcg gtggagaaaa catcatcata ataataataa agccgaataa 180
ctttctccac acgtagtgtg tgtcaataaa tcggttgttt ctcgcagagt ttgaaaggaa 240
acttttagag ttcttaaaca aagaatgtgg ttttggaacc actgtgccgc atcctgtctg 300
ccatgaatga tgcgtccttt cctgccagca cataaagccc cggccagttg catctattcc 360
acacactgat agacaatgtt tgtgcactga tttatactca gccagctatt aaataagaat 420
tttcgttatt acgctgccgg ggattttatt atgtgacttt caaaagcccc cgccaacttg 480
ttgtggcttt ggcttaaagc cgcgccaagc tgaaaattgt tattttaacc ttttaaagtc 540
agctttatca gcaccaaatg cagggccatt atgcgtggcc cacgcggcgt atgagcaacg 600
ttaaatctgc gaataattca aagacgtgcg tttcttttta cagccgcctt gtcggtaatc 660
ccattaagga gttatctggc gaaacatttt tgcacaacac tcgactggag gcactgtgag 720
tatactttgt ttgttttaat ttaaaaaatc acaaaatata tcatgaagtt gttaaaacgg 780
gattgaattt gatttgaata tatgatttca tttaatttct caccaatatt taagttcgtt 840
ggccctaatg ccgattcaca tcagcagcag tttaatggag ccactgaata tatcgttttt 900
gaatctcacg ggaattcggt atgatcacat cgacttcgag gccatcaatt caatgcgcaa 960
cctgacgtac atcatctacg atcgattctt ctactgctca atgacgccgc gggtgagaat 1020
gtgcaagccg tccaccgatg gagtcagctc ctttcaagat cttctgagca aaccagtgtt 1080
acgttactcc gcctgggtga tggccacgtt gacaatagct ggaaatgtgc tagtcctatg 1140
gggacgcttt atctaccgcg atgagaatgt ggctgtgaca atggtgattc ggaatctggc 1200
tctggctgat atgctgatgg gcttctattt ggtcaccatt ggagtgcagg actatcgcta 1260
ccgcaacgag tattacaagt aagtcaaatg tgatggattg gcatttgtta cgccccaatc 1320
ttttagggtt gtgctggact ggataacctc gtggcagtgc acactgattg gaacgctggc 1380
cgtgagctcc tcggaggttt cgatgcttat cctggctttt atgtcgctgg agcggtttct 1440
cttgatcgcc gatccctttc gcggccatcg cagcattgga aatcgggtga tgtggctggc 1500
tttgatttgc atttggatta cgggcgtggg cttggctgtg gcaccggtgc tcctgtggcg 1560
gaccagtacc ctaccctact acggatccta ttccggaacc tgtgagcaca atgtttatga 1620
taaaatgtag tttaccaagt gacgaaatta aacttaacag gctttccatt gcacatccac 1680
gaggccttcc cgatgggctg gctttactcg gctttcgtgt tcctgggcgt caatctgctg 1740
ctgctggtga tgattgcgat gctctacacg gcactcctta tttcgatatg gcgaacgagg 1800
agcgccactc cgctcactct gttggattgc gaattcgctg tgcgattctt tttcatcgtg 1860
```

ctgaccgact tcctgtgctg ggtgcccatc atcgttatga agatctgggt ttttttcaac 1920
tacaatatct ccgatgacat ctacgcctgg ttggtggtct ttgtgctgcc acttaactcg 1980
gcggttaatc ctctcttgta caccttcacc acgcccaagt accgaaatca aatcttcctt 2040
cgcggctgga agaagataac ttctagaaag cgagcggagg ctggcaatgg aaatgtggcc 2100
accacaacga cgggcacggc aactggatcc tctcagcatc cggatgactt cacgatcttt 2160
gccaaagccg cgatgagatg ccattaggtc ctaacgctgt cgaactgaag gattcttgtg 2220
gcgcagatcg aacttgtata ccaccagatt tagcagcatt tattttgtct atgatatcga 2280
aaatctatat atgtatgtga atatctatcc ctgtccgtgt aagataacca aatgaaagtg 2340
tgaaataaat tgaatgtgat atttgtaagg caattttacg gaatttaaaa tgggatgcga 2400
gttagtgtaa gattaagcta ttatctaagt tttggtaata gttttaaata ctttttattt 2460
agaaatagca gcgctatcat ttttggcatg gtgtgattct ttctggtccg agtaaatctt 2520
tgccattacc aacttggagc gcacgatgtg aaactttcgc atttcctcca ggtcctccat 2580
ctgcaccacc atgttatgct tacgctccag gtgcatgcgt agtctaaaaa attattaagg 2640
agtgttactt tcaaataatg tacactttaa acaaaataac taacttgtta cgagctggat 2700
agctctgatg gcagtgggcg cactcaaatg gcatggccgt atgcgtgcac tcgtgattcc 2760
tgagagatga agatgaggta aaggacttct ggcagccggg aaactcgcaa atgaagcgag 2820
atccagagtg cttaacctgc tggtggacag tcaaatggga tcgctgtgcg aaacgacgat 2880
cacagagatc acaggaatag ggaaagtctc tgttatgata gcgtttatga atctccaggg 2940
ctttggcaat ggcaaaaacc ttgccgcatg tgtcgcagtg gactccagac tgcctgtgct 3000
tcgaatcgac gtgaactttc agcgcctgct ttgtgaggta aacccggtcg cactggctgc 3060
aacgaaagtt gggacgattg gccaaatgaa tcatctggtg cttggccatg cgggaatggt 3120
tcctgaactc cttgtggcag aacttgcagt tggattttcc cgtatgccga ttgcgtggca 3180
caatgtc 3187

<210> 26
<211> 1080
<212> DNA
<213> Drosophila

<400> 26
atgacgccgc gggtgagaat gtgcaagccg tccaccgatg gagtcagctc ctttcaagat 60
cttctgagca aaccagtgtt acgttactcc gcctgggtga tggccacgtt gacaatagct 120
ggaaatgtgc tagtcctatg ggacgcttt atctaccgcg atgagaatgt ggctgtgaca 180
atggtgattc ggaatctggc tctggctgat atgctgatgg gcttctattt ggtcaccatt 240
ggagtgcagg actatcgcta ccgcaacgag tattacaagg ttgtgctgga ctggataacc 300
tcgtggcagt gcacactgat tggaacgctg gccgtgagct cctcggaggt ttcgatgctt 360
atcctggctt ttatgtcgct ggagcggttt ctcttgatcg ccgatccctt tcgcggccat 420
cgcagcattg gaaatcgggt gatgtggctg gctttgattt gcatttggat tacgggcgtg 480
ggcttggctg tggcaccggt gctcctgtgg cggaccagta ccctacccta ctacggatcc 540
tattccggaa cctgctttcc attgcacatc cacgaggcct tcccgatggg ctggctttac 600
tcggctttcg tgttcctggg cgtcaatctg ctgctgctgg tgatgattgc gatgctctac 660
acggcactcc ttatttcgat atggcgaacg aggagcgcca ctccgctcac tctgttggat 720
tgcgaattcg ctgtgcgatt ctttttcatc gtgctgaccg acttcctgtg ctgggtgccc 780
atcatcgtta tgaagatctg ggtttttttc aactacaata tctccgatga catctacgcc 840
tggttggtgg tctttgtgct gccacttaac tcggcggtta atcctctctt gtacaccttc 900
accacgccca agtaccgaaa tcaaatcttc cttcgcggct ggaagaagat aacttctaga 960
aagcgagcgg aggctggcaa tggaaatgtg gccaccacaa cgacgggcac ggcaactgga 1020
tcctctcagc atccggatga cttcacgatc tttgccaaag ccgcgatgag atgccattag 1080

<210> 27
<211> 359
<212> PRT
<213> Drosophila

<400> 27
Met Thr Pro Arg Val Arg Met Cys Lys Pro Ser Thr Asp Gly Val Ser
1 5 10 15
Ser Phe Gln Asp Leu Leu Ser Lys Pro Val Leu Arg Tyr Ser Ala Trp
20 25 30

```
Val Met Ala Thr Leu Thr Ile Ala Gly Asn Val Leu Val Leu Trp Gly
        35                  40                  45
Arg Phe Ile Tyr Arg Asp Glu Asn Val Ala Val Thr Met Val Ile Arg
    50                  55                  60
Asn Leu Ala Leu Ala Asp Met Leu Met Gly Phe Tyr Leu Val Thr Ile
65                  70                  75                  80
Gly Val Gln Asp Tyr Arg Tyr Arg Asn Glu Tyr Tyr Lys Val Val Leu
                85                  90                  95
Asp Trp Ile Thr Ser Trp Gln Cys Thr Leu Ile Gly Thr Leu Ala Val
            100                 105                 110
Ser Ser Ser Glu Val Ser Met Leu Ile Leu Ala Phe Met Ser Leu Glu
        115                 120                 125
Arg Phe Leu Leu Ile Ala Asp Pro Phe Arg Gly His Arg Ser Ile Gly
    130                 135                 140
Asn Arg Val Met Trp Leu Ala Leu Ile Cys Ile Trp Ile Thr Gly Val
145                 150                 155                 160
Gly Leu Ala Val Ala Pro Val Leu Leu Trp Arg Thr Ser Thr Leu Pro
                165                 170                 175
Tyr Tyr Gly Ser Tyr Ser Gly Thr Cys Phe Pro Leu His Ile His Glu
            180                 185                 190
Ala Phe Pro Met Gly Trp Leu Tyr Ser Ala Phe Val Phe Leu Gly Val
        195                 200                 205
Asn Leu Leu Leu Leu Val Met Ile Ala Met Leu Tyr Thr Ala Leu Leu
    210                 215                 220
Ile Ser Ile Trp Arg Thr Arg Ser Ala Thr Pro Leu Thr Leu Leu Asp
225                 230                 235                 240
Cys Glu Phe Ala Val Arg Phe Phe Phe Ile Val Leu Thr Asp Phe Leu
                245                 250                 255
Cys Trp Val Pro Ile Ile Val Met Lys Ile Trp Val Phe Phe Asn Tyr
            260                 265                 270
Asn Ile Ser Asp Asp Ile Tyr Ala Trp Leu Val Val Phe Val Leu Pro
        275                 280                 285
Leu Asn Ser Ala Val Asn Pro Leu Leu Tyr Thr Phe Thr Thr Pro Lys
    290                 295                 300
Tyr Arg Asn Gln Ile Phe Leu Arg Gly Trp Lys Lys Ile Thr Ser Arg
305                 310                 315                 320
Lys Arg Ala Glu Ala Gly Asn Gly Asn Val Ala Thr Thr Thr Thr Gly
                325                 330                 335
Thr Ala Thr Gly Ser Ser Gln His Pro Asp Asp Phe Thr Ile Phe Ala
            340                 345                 350
Lys Ala Ala Met Arg Cys His
        355
```

<210> 28
<211> 3759
<212> DNA
<213> Drosophila

<400> 28

```
cactcagatt tgccgacagc agaaatattt tagggttgta tcttagaggt ggagaaacat 60
cgatgccaac atcgatgtca tcgatgtttt cgaaaaatca aaaacgtcga tgtttttcga 120
ggaatcgtcg atgttttcag cgcatagaaa cagcaaagtg tgtaaattgt aagtctaaat 180
taatttccat attttctaat acaattatta atagcatatt gcatatttat tgtgactcgc 240
cggtggtgta ggctggtgtc ttggaagaca tggataaata tctgaaatgt aagttttttt 300
tttaaagatc aatttatttg ttttttgcca aagacatcga aaacatggat gttcgccaac 360
atcgatgttt ctccacttct aatgccaacg aaaacgtgac tgggaacgta tcgatgggaa 420
aaacgttaca aaaataccaa ctatcgcata gcggaatcac cttggcgggc tattcgaatg 480
tatctcacaa atagttttta aggaacctac agtcagtacc aaatacttaa aattccattt 540
atattttatt ttatgctaca gaaaactgat aagctgcatt acttttgaaa agttaaaaat 600
```

tgtgtttttt aaatgattta tttatataaa taagcttatg taatactatt gttaatatgt 660
aaaaaaaaaa gaataaaaga aaaagaatag tttacctaag cagttgcttg aataatacat 720
acataattta agttttatac ataatttatc aaccctcgct cagtaggaaa aaaagatggc 780
tcagaaatga gcaacgaatt ggaatgtccg ataagccttt tttatcaaaa taataatttt 840
tacattccct tttcataacg ctttcccagt tagtttgtgc caatcgtaaa acagcatccg 900
caaagtgaag tcgctgctat tttgtagtac cactacattc ttatttattt aagtaatatt 960
taaagaatgt ccagctaact atatgattgc tagttcaaaa atgttacttt ctgcatcgat 1020
cctaatatat ttcttgctga atctgcaatc ttcgagtgct gagattgccg actgcagttt 1080
ctatgatacc gttgatatct cggaaggcca gaggctctcg aatggatcat acctctacga 1140
gggattactt attcccgccc atttgacggc caagtatgaa ttcaagcttc tggcgaatgg 1200
agataaggag caggtgccaa gtcacgtgag aggatgtgtg tgcaagctaa ggacttgcgt 1260
caggttttgc tgcccacatg accacataat ggatatgggc gaatgctacg ccaacatgac 1320
gacggaggag aacgaattac tggaccccat gctaaatgtg accctggacg acggatcagt 1380
ggtgcaaagg cactataaga aggagctcat ggtgcagtgg gatctgccaa agccctgcga 1440
tgatatgttt tacctagaca accgggatat aatggatgag tacaccttgt tcgaggtatg 1500
gctgctcgaa aaataagcat ttttgggcag gatttattta tgctttggct ctctgcagaa 1560
tggaagattg ctgcgccact acgatcaagt gtatttggac aagtcggagt actgtctgca 1620
gcaccgaaca ttcggcgaag gtaacaacaa ctccattcgc attataсссс acaactgctt 1680
gatactgcca tccagaacgg gacagaccgt tggtaagatg atgcccattg cttgcatggt 1740
agctcccagt tctatatctt tcctatttcc atagtgatga tcacatcgtt gatatgcttg 1800
gttcttacca tcgccgtgta cctgtgcgtt aagaagctaa tgaatttgga gggcaagtgt 1860
ttcatctgtt acatgatgtg cctatttttc ggatacctgt tcctactcct tgatctgtgg 1920
gagttatcgc tcgatttctg taaagcagca ggttagctta gaggggatgc atttcaacac 1980
aacttgacaa accattgacc atttacgaca ggtttcttgg gatacttttt cgtcatggct 2040
gcgttctttt ggctatctat catcagtcgg cactattgga aatgtcttac caatccctgt 2100
gcctcgatga acattcggtc agaacgtgcg tttctgctat acagctgctt cgcctgggca 2160
atgcctttgg ccttgaccgg agtcacctat ctggctgata acgttgtgaa caacgaggag 2220
tggcagcctc gcgtgggcga cgaagggcat tgctggatct atagtgggtt ctgttttgta 2280
tccggcttgt ggtaccagct caccaatgtt tttctttcca ccctcagcta aaagctggtc 2340
tgcaatggtg tacttctatg gaccgatggt gctcctgatt ctcttcaaca taactatgtt 2400
tgttctgacg gctaaacaca taatagattc caagaggact cttcgaaaga ttgctcgcaa 2460
tgaaggtaga atacagaagc tgaattcgga caaacagaag tatatcaatt ataatagttg 2520
aaaatgtcag tttaagataa atattatcaa ttcgttacag ttacacacaa tttctgctgc 2580
tcttcactgt gatggggatg tcgtggagct ttgagatatt ctcctacttg gtgcaacgcg 2640
agaaactgtg ggtcaatatc tttctggtag ctgactactt caactggtcc caaggcgtca 2700
ttatatttgt gctattcatt ctgcggcgca aaactcttgt actattcaag aaacagtaag 2760
tcttatagtt catagttcgg gtaaccagtt ccctaaaatg tctcttttca gaatatttcc 2820
caaacagaga gcctttagta gatctgcaac tcaaagtacg attgaatcca tttctcaaac 2880
aaaaaggcat tttaacatga catagtattt tcttctcttc gaattaatca atttaatagt 2940
aaagtataac tttttttctc acaaaatgta tgaatatatt tcacttcatt ctacattcta 3000
tatagaattc atttaatcac aaaatgtatg aatatttttc tttatatgac gtgcatcgac 3060
tatagaaaat cgactatagt gtgccacaat atattttact catcatttcc taatcaaact 3120
attgcaaact taactttgag ccgaggaaaa accctgtatg tctatcaaaa ctatgagtca 3180
aaagtatgac gcgatgctga aagtagttgc ctttaaatat tgaagtgata aatgtagaaa 3240
gaagacattt gtatggaaat aaactaagtt tgtaaatcta cattcgcttt atttcatgta 3300
caaattttaa tttaaaaact ttgaaggtaa atagtgcttt atcagatgcg gattacacgg 3360
tagcagttat catcgatgtc gaatcttctt gtttatgtga ccgcttttct tcgctgtcgc 3420
actggagttc gtggatagtt tgctctttgt attcgaagtg ggtgtgccga ggagatcctt 3480
gttccctcca ctgttcgtag atccagacac cattccgctc actacgctgc cctcattgcg 3540
atgcgagggt gaatctgtaa ataaaaaggt gtaattaaac atatatactt ttctggagaa 3600
aagcgctaaa gttgaatctc aatgtgctgt gctgccttac ccttttcgtc tttcggtgtg 3660
ctcttgattg tcgtggttcc gtcgtgggcc gttgtggtgt tctgattgat tgagatctga 3720
ttgtcgctgg agtcttcatt tgtgatggaa tccgattcg 3759

<210> 29
<211> 1449
<212> DNA
<213> Drosophila

<400> 29

```
atgttacttt ctgcatcgat cctaatatat ttcttgctga atctgcaatc ttcgagtgct 60
gagattgccg actgcagttt ctatgatacc gttgatatct cggaaggcca gaggctctcg 120
aatggatcat acctctacga gggattactt attcccgccc atttgacggc caagtatgaa 180
ttcaagcttc tggcgaatgg agataaggag caggtgccaa gtcacgtgag aggatgtgtg 240
tgcaagctaa ggacttgcgt caggttttgc tgcccacatg accacataat ggatatgggc 300
gaatgctacg ccaacatgac gacggaggag aacgaattac tggaccccat gctaaatgtg 360
accctggacg acggatcagt ggtgcaaagg cactataaga aggagctcat ggtgcagtgg 420
gatctgccaa agccctgcga tgatatgttt tacctagaca accgggatat aatggatgag 480
tacaccttgt tcgagaatgg aagattgctg cgccactacg atcaagtgta tttggacaag 540
tcggagtact gtctgcagca ccgaacattc ggcgaaggta acaacaactc cattcgcatt 600
ataccccaca actgcttgat actgccatcc agaacgggac agaccgttgt gatgatcaca 660
tcgttgatat gcttggttct taccatcgcc gtgtacctgt gcgttaagaa gctaatgaat 720
ttggagggca agtgtttcat ctgttacatg atgtgcctat ttttcggata cctgttccta 780
ctccttgatc tgtgggagtt atcgctcgat ttctgtaaag cagcaggttt cttgggatac 840
tttttcgtca tggctgcgtt cttttggcta tctatcatca gtcggcacta ttggaaatgt 900
cttaccaatc cctgtgcctc gatgaacatt cggtcagaac gtgcgtttct gctatacagc 960
tgcttcgcct gggcaatgcc tttggccttg accggagtca cctatctggc tgataacgtt 1020
gtgaacaacg aggagtggca gcctcgcgtg ggcgacgaag ggcattgctg gatctatact 1080
aaaagctggt ctgcaatggt gtacttctat ggaccgatgg tgctcctgat tctcttcaac 1140
ataactatgt ttgttctgac ggctaaacac ataatagatt ccaagaggac tcttcgaaag 1200
attgctcgca atgaaggtag aatacagaag ctgaattcgg acaaacagaa ttacacacaa 1260
tttctgctgc tcttcactgt gatggggatg tcgtggagct ttgagatatt ctcctacttg 1320
gtgcaacgcg agaaactgtg ggtcaatatc tttctggtag ctgactactt caactggtcc 1380
caaggcgtca ttatatttgt gctattcatt ctgcggcgca aaactcttgt actattcaag 1440
aaacagtaa                                                         1449
```

```
<210> 30
<211> 482
<212> PRT
<213> Drosophila

<400> 30
Met Leu Leu Ser Ala Ser Ile Leu Ile Tyr Phe Leu Leu Asn Leu Gln
 1               5                  10                  15
Ser Ser Ser Ala Glu Ile Ala Asp Cys Ser Phe Tyr Asp Thr Val Asp
            20                  25                  30
Ile Ser Glu Gly Gln Arg Leu Ser Asn Gly Ser Tyr Leu Tyr Glu Gly
        35                  40                  45
Leu Leu Ile Pro Ala His Leu Thr Ala Lys Tyr Glu Phe Lys Leu Leu
    50                  55                  60
Ala Asn Gly Asp Lys Glu Gln Val Pro Ser His Val Arg Gly Cys Val
65                  70                  75                  80
Cys Lys Leu Arg Thr Cys Val Arg Phe Cys Cys Pro His Asp His Ile
                85                  90                  95
Met Asp Met Gly Glu Cys Tyr Ala Asn Met Thr Thr Glu Glu Asn Glu
            100                 105                 110
Leu Leu Asp Pro Met Leu Asn Val Thr Leu Asp Asp Gly Ser Val Val
        115                 120                 125
Gln Arg His Tyr Lys Lys Glu Leu Met Val Gln Trp Asp Leu Pro Lys
    130                 135                 140
Pro Cys Asp Asp Met Phe Tyr Leu Asp Asn Arg Asp Ile Met Asp Glu
145                 150                 155                 160
Tyr Thr Leu Phe Glu Asn Gly Arg Leu Leu Arg His Tyr Asp Gln Val
                165                 170                 175
Tyr Leu Asp Lys Ser Glu Tyr Cys Leu Gln His Arg Thr Phe Gly Glu
            180                 185                 190
Gly Asn Asn Asn Ser Ile Arg Ile Ile Pro His Asn Cys Leu Ile Leu
        195                 200                 205
Pro Ser Arg Thr Gly Gln Thr Val Val Met Ile Thr Ser Leu Ile Cys
    210                 215                 220
```

```
Leu Val Leu Thr Ile Ala Val Tyr Leu Cys Val Lys Lys Leu Met Asn
225                 230                 235                 240
Leu Glu Gly Lys Cys Phe Ile Cys Tyr Met Met Cys Leu Phe Phe Gly
                245                 250                 255
Tyr Leu Phe Leu Leu Leu Asp Leu Trp Glu Leu Ser Leu Asp Phe Cys
            260                 265                 270
Lys Ala Ala Gly Phe Leu Gly Tyr Phe Phe Val Met Ala Ala Phe Phe
        275                 280                 285
Trp Leu Ser Ile Ile Ser Arg His Tyr Trp Lys Cys Leu Thr Asn Pro
    290                 295                 300
Cys Ala Ser Met Asn Ile Arg Ser Glu Arg Ala Phe Leu Leu Tyr Ser
305                 310                 315                 320
Cys Phe Ala Trp Ala Met Pro Leu Ala Leu Thr Gly Val Thr Tyr Leu
                325                 330                 335
Ala Asp Asn Val Val Asn Asn Glu Glu Trp Gln Pro Arg Val Gly Asp
            340                 345                 350
Glu Gly His Cys Trp Ile Tyr Thr Lys Ser Trp Ser Ala Met Val Tyr
        355                 360                 365
Phe Tyr Gly Pro Met Val Leu Leu Ile Leu Phe Asn Ile Thr Met Phe
    370                 375                 380
Val Leu Thr Ala Lys His Ile Ile Asp Ser Lys Arg Thr Leu Arg Lys
385                 390                 395                 400
Ile Ala Arg Asn Glu Gly Arg Ile Gln Lys Leu Asn Ser Asp Lys Gln
                405                 410                 415
Asn Tyr Thr Gln Phe Leu Leu Leu Phe Thr Val Met Gly Met Ser Trp
            420                 425                 430
Ser Phe Glu Ile Phe Ser Tyr Leu Val Gln Arg Glu Lys Leu Trp Val
        435                 440                 445
Asn Ile Phe Leu Val Ala Asp Tyr Phe Asn Trp Ser Gln Gly Val Ile
    450                 455                 460
Ile Phe Val Leu Phe Ile Leu Arg Arg Lys Thr Leu Val Leu Phe Lys
465                 470                 475                 480
Lys Gln
```

<210> 31
<211> 4469
<212> DNA
<213> Drosophila

<400> 31

```
acgttgtcag tgtgtgtgac ggttgttaat ttcatttgac ttgcgagggt tttaatattg 60
caatggaaat tcggtacaat ttgccgtcca ttaaataagc acgcaattaa ttgtggcagc 120
catgacaggc gtggcttgcc accaattaag tggaatacta tatggctata tcgctatatc 180
gctatatgtc tatatggcat atagatgccc acagcccctt ggggtcgatc gcctacgccg 240
tctataatcc ccagctaaac ttcgagacat cgcatccggc accgcaatcc gaaatcggaa 300
tgctgatgag gcaacagagt ggacgcggcc ggaaacaagg cactcctccc atcaatctgt 360
cagacgggcc agtgagtcct ggatgtcctg gtgaacgcag ctcagctccc agcgggtgag 420
gccagttgcc cagacagtta gtccgccagg cagtcagtcg gtctgtcagt gagtcagtta 480
gtcagtgagc gagtgagtca gtcagtcaat cgaacaatca atcatacgga atgacagcta 540
cgtcggctta agtcaaaggc agtcggccaa gaagggatct tgtctaactt acagtcgcaa 600
aattcaatag cctgtggcaa tcgcacttgc aactttaata attccaagaa cgaagcagtc 660
ataatgttta ctgaaataat gttatgccac aattggcaca tacaagctca taatagtgtt 720
ttattctgtt tttaattgca ctttaaaatt gatgattgta ttagccatca taggcatggc 780
ttgagctgtt catttagcca tttgattttg tctacgttgg ttttttatca agtttcttgc 840
cgtgtaacct tttttttttt tttgggctct cccaacaatt tgagctcctt tttcgttttc 900
gcattccctt aaacatttgc tggggaaaca cattgattta tatgccggca tcgccaaaaa 960
ttgaatgtta ttgcttgaat ggccccaatt tacttaaaag ctacgctcct ccgccccact 1020
gcccctctgc cgctccactg acccctgacc ccattataca tgttgtgggc tccactggct 1080
```

cctgattggt atttatgagc gtgtttattt attatatgac ccaacacttg ataaaattaa 1140
acggaaatgg ttggcaattt tgcatttgat tctgacaccc aaacacacag gcaggcaaat 1200
tcgcaaatga actcatctgc ccaatattcc gaccataaag catcccgacc aacggcactc 1260
acatataagt tagttaagta tgtatattag tggcttggca acttaccaga aagtagacag 1320
caaatggcat cacgaccacg gcgaccagga gatcggcgat ggccaacgag actataaaat 1380
aattcgtgac cgtctgcagg gagcgctcgc ggcacacgga caggatgacc aggatgttgc 1440
cgaagagggt caggatgggg aacagaatca ggatgagcgc ccagtagttg tggtccactg 1500
aaacgaagcg agttgtagtt acattatatc cgggatgcat ccatccgcag gctctcctgc 1560
tgctgccacc caaattccct ggccaaagtt caccttgcag gaagatttca gggaaagctc 1620
aacgggcacg taagtcaatt acacaggcag tccgtttaat caaggactct ggctgacaac 1680
taattttaat ggcaacacaa gtaaggacta gcacacaaag acgtgaatct taaaggctgt 1740
tttaatataa tctatctaat atagagtgat gcatgttacc aacagctgcg ttgttaaaaa 1800
ttaaagttat ttcacatatt aaaaatggca aaatatggct atgtaatatg aaagttatta 1860
aacttgaaat acctttatct tcattatgaa acaaattaaa caaacaaata tgtaggaact 1920
tttcttgaga ttgcacgaat aaatgagtta tgtaacgcgt ttaaattaaa ttcttgctat 1980
caaaactgat ttctaaatga acatctttat aaataaatta attatcggcg gtttcattca 2040
aatactccca tcttagaaat ctaaggaatt ctgttctcaa aacctagtca cctcactatc 2100
cgatgaaaat atattatact aaagaagttg ctgttgcaag gtttaaatga tgctccaact 2160
tgatgataat gtttaaaggt atttaatgag attcaaggct atctaatcta atttaataat 2220
tttaaagcct ttgaatgata ttcaataata tctaatgagc tttaatgatt tttaaagatt 2280
tataatgcgg ttttaagatg ttaagttgta gtagtttgat gatttcgaat ggtaatcctg 2340
cctcgtactc acccactctc agctcgccgc aggagtcggt taggttgagc gtctcgttgg 2400
tgcagttcag ccggtagccg tcgaggtaca gtcccgtcag gttctgctcg aggagcagtc 2460
cgctgcccct gctccccgtg atggccactc ccggcgcaat ggtgcttgtt cccgaaaccg 2520
tgctcccatt gtagttggga aaataatcgt taaagttttc gagcagcagc atgaaggtgc 2580
tgccccctga tcctccagcg ccgcctacac ctccgcccgc cgtaccggcc agtccgccag 2640
ccacgcccac atccccagag ccgcctccat cgtaactggt cgtcgagggc gccgttcgat 2700
tgcccagcgt ctcgcgcgcc accgttgctg tccgctggag cagttccgca tcaccgaagc 2760
cgcccagcat gccatagtcc gtttccgcat cagccacatt gtccagggct gcctcctccc 2820
cgctgtcgat cagcgcgatg ccagcatcgc cgccgtcgcc cgctcctgct gcccctccgc 2880
ccgccagctg ttgttccaca aagctgcgca gcgagtaacg cccccgctca ccatcctccg 2940
cactggagtt gctctgggct ggctgcgtgc tgttcagtgc ctccagccag gccagggttc 3000
cgttggcgcc cgtcaccgcc aagtttgtgc tgttctcggg cccgcccact gccgcctcac 3060
tgcccaccga acccaagagc cattcgctga tgtcgtcgat cagcgcgtag ccggcgtcct 3120
ccgccgccgt ttccacatcc tcgtagagcg gcaagtagcc ggctggcatt gtggggccct 3180
caacgcccgc ctgcgcatcg ctgccgttcg ccggttgaga gctgtccgct tccaccgcca 3240
ggtggtgatt tatccacgtg gctcccgcgg atggggctgc cgtggtcgag gtggaagcag 3300
cggctgtgct caaggtcgcc gatgtggttg ctgccgttgc tgcagttgcc gcagttgctc 3360
ccgttgccgc cgcagttgtt gccgccgtgc tgctcagtgg ctgggctgcc atggtgccgc 3420
cggttccgct gctgctgatg ccccgccgcc aatcgaatgg gctcaacatc tagtggggcg 3480
ggggcgtggc agcggaggcg agctgcaact cctgcaacgc cagcggcatg tttccgtttc 3540
ctgcagcgaa agagacgaag ttcggaggtg agaaaacata aatcacaaaa atttccatta 3600
taatcaaagt tcttgaagat agttcttaac gatacgcaca actaagatta caattttaag 3660
cctaattggt cgcgttaatg ttgcttaatt tgataaattc ttgtaatccg caaatgagag 3720
tcatggttct tttcgactta atatttttag gatatctaaa aaacaaagtc gttctcaatt 3780
tttcctactt acacattaca cattgcatgc cccttcgttt gcattttcaa cacaaggcgt 3840
acttcatgct tgaccacgcc cccccaacca acacaacaac ataccatcca gcaccacgcc 3900
cccatgccat tttggacgta cggacccttt gcccaactgc catggccgca tgtttaatta 3960
actctttaat caatttattg tcaattattc tacaaaaagc gacagcgaaa attgtgtaat 4020
ttaattaaat tagcaaaagg tttttgtgtc tctctcgccc caccaaccat tccgcccact 4080
cccacttcca ttctcgcaac caaaatgcca cgcccccatg ctcgacgctg tcttgctttt 4140
acacattttg caattttcgc ttttgtattt gccaattggc gctgattaaa gccgcccgag 4200
tcgatgggcc ataattaaaa gatggcctaa tgagaggcga ggcgagtacg tctcgagttg 4260
gctgggggtc aaaggtaatc ggatcggagt ggtcttaagt cccctgcctt ctgcctctgt 4320
cttccggaga tcgggttatt cgaactctga gcttgtggat tctggactct gaattctgga 4380
ttccggatcg gaggcagaga tccgagagct gcacacccag acggctggca ctaatccaat 4440
ttgggcacac ttccaggcga cgagaggca 4469

<210> 32
<211> 1389

<212> DNA
<213> Drosophila

<400> 32
atgttgagcc cattcgattg gcggcggggc atcagcagca gcggaaccgg cggcaccatg 60
gcagcccagc cactgagcag cacggcggca acaactgcgg cggcaacggg agcaactgcg 120
gcaactgcag caacggcagc aaccacatcg gcgaccttga gcacagccgc tgcttccacc 180
tcgaccacgg cagccccatc cgcgggagcc acgtggataa atcaccacct ggcggtggaa 240
gcggacagct ctcaaccggc gaacggcagc gatgcgcagg cgggcgttga gggccccaca 300
atgccagccg gctacttgcc gctctacgag gatgtggaaa cggcggcgga ggacgccggc 360
tacgcgctga tcgacgacat cagcgaatgg ctcttgggtt cggtgggcag tgaggcggca 420
gtgggcgggc ccgagaacag cacaaacttg gcggtgacgg gcgccaacgg aaccctggcc 480
tggctggagg cactgaacag cacgcagcca gcccagagca actccagtgc ggaggatggt 540
gagcgggggc gttactcgct gcgcagcttt gtggaacaac agctggcggg cggagggggca 600
gcaggagcgg gcgacggcgg cgatgctggc atcgcgctga tcgacagcgg ggaggaggca 660
gccctggaca atgtggctga tgcggaaacg gactatggca tgctgggcgg cttcggtgat 720
gcggaactgc tccagcggac agcaacggtg gcgcgcgaga cgctgggcaa tcgaacggcg 780
ccctcgacga ccagttacga tggaggcggc tctggggatg tgggcgtggc tggcggactg 840
gccggtacgg cgggcggagg tgtaggcggc gctggaggat cagggggcag caccttcatg 900
ctgctgctcg aaaactttaa cgattatttt cccaactaca atgggagcac ggtttcggga 960
acaagcacca ttgcgccggg agtggccatc acggggagca ggggcagcgg actgctcctc 1020
gagcagaacc tgacgggact gtacctcgac ggctaccggc tgaactgcac caacgagacg 1080
ctcaacctaa ccgactcctg cggcgagctg agagtggtgg accacaacta ctgggcgctc 1140
atcctgattc tgttccccat cctgaccctc ttcggcaaca tcctggtcat cctgtccgtg 1200
tgccgcgagc gctccctgca gacggtcacg aattatttta tagtctcgtt ggccatcgcc 1260
gatctcctgg tcgccgtggt cgtgatgcca tttgctgtct actttctgga gccagtggag 1320
cccacaacat gtataatggg gtcaggggtc agtggagcgg cagaggggca gtggggcgga 1380
ggagcgtag 1389

<210> 33
<211> 462
<212> PRT
<213> Drosophila

<400> 33
Met Leu Ser Pro Phe Asp Trp Arg Arg Gly Ile Ser Ser Ser Gly Thr
1 5 10 15
Gly Gly Thr Met Ala Ala Gln Pro Leu Ser Ser Thr Ala Ala Thr Thr
20 25 30
Ala Ala Ala Thr Gly Ala Thr Ala Ala Thr Ala Ala Thr Ala Ala Thr
35 40 45
Thr Ser Ala Thr Leu Ser Thr Ala Ala Ala Ser Thr Ser Thr Thr Ala
50 55 60
Ala Pro Ser Ala Gly Ala Thr Trp Ile Asn His His Leu Ala Val Glu
65 70 75 80
Ala Asp Ser Ser Gln Pro Ala Asn Gly Ser Asp Ala Gln Ala Gly Val
85 90 95
Glu Gly Pro Thr Met Pro Ala Gly Tyr Leu Pro Leu Tyr Glu Asp Val
100 105 110
Glu Thr Ala Ala Glu Asp Ala Gly Tyr Ala Leu Ile Asp Asp Ile Ser
115 120 125
Glu Trp Leu Leu Gly Ser Val Gly Ser Glu Ala Ala Val Gly Gly Pro
130 135 140
Glu Asn Ser Thr Asn Leu Ala Val Thr Gly Ala Asn Gly Thr Leu Ala
145 150 155 160
Trp Leu Glu Ala Leu Asn Ser Thr Gln Pro Ala Gln Ser Asn Ser Ser
165 170 175
Ala Glu Asp Gly Glu Arg Gly Arg Tyr Ser Leu Arg Ser Phe Val Glu
180 185 190
Gln Gln Leu Ala Gly Gly Gly Ala Ala Gly Ala Gly Asp Gly Gly Asp

195 200 205
Ala Gly Ile Ala Leu Ile Asp Ser Gly Glu Glu Ala Ala Leu Asp Asn
210 215 220
Val Ala Asp Ala Glu Thr Asp Tyr Gly Met Leu Gly Gly Phe Gly Asp
225 230 235 240
Ala Glu Leu Leu Gln Arg Thr Ala Thr Val Ala Arg Glu Thr Leu Gly
245 250 255
Asn Arg Thr Ala Pro Ser Thr Thr Ser Tyr Asp Gly Gly Gly Ser Gly
260 265 270
Asp Val Gly Val Ala Gly Gly Leu Ala Gly Thr Ala Gly Gly Gly Val
275 280 285
Gly Gly Ala Gly Gly Ser Gly Gly Ser Thr Phe Met Leu Leu Leu Glu
290 295 300
Asn Phe Asn Asp Tyr Phe Pro Asn Tyr Asn Gly Ser Thr Val Ser Gly
305 310 315 320
Thr Ser Thr Ile Ala Pro Gly Val Ala Ile Thr Gly Ser Arg Gly Ser
325 330 335
Gly Leu Leu Leu Glu Gln Asn Leu Thr Gly Leu Tyr Leu Asp Gly Tyr
340 345 350
Arg Leu Asn Cys Thr Asn Glu Thr Leu Asn Leu Thr Asp Ser Cys Gly
355 360 365
Glu Leu Arg Val Val Asp His Asn Tyr Trp Ala Leu Ile Leu Ile Leu
370 375 380
Phe Pro Ile Leu Thr Leu Phe Gly Asn Ile Leu Val Ile Leu Ser Val
385 390 395 400
Cys Arg Glu Arg Ser Leu Gln Thr Val Thr Asn Tyr Phe Ile Val Ser
405 410 415
Leu Ala Ile Ala Asp Leu Leu Val Ala Val Val Val Met Pro Phe Ala
420 425 430
Val Tyr Phe Leu Glu Pro Val Glu Pro Thr Thr Cys Ile Met Gly Ser
435 440 445
Gly Val Ser Gly Ala Ala Glu Gly Gln Trp Gly Gly Gly Ala
450 455 460

<210> 34
<211> 11501
<212> DNA
<213> Drosophila

<400> 34
tttgtaccgc agtcaagcgc aaaaattcct cgaaaaacca tagaccaagg tgtttatgag 60
tttattaata ctgtgtaatt attatttgtg atatatgtgt gggaaattgg taatagatca 120
aacatgaaag taatggtatc tgtttttctg tattaaaata acatatctga taatttccta 180
aaggatgtta atgttaattt aataggtatt ccaaaactaa gagtgaactg gcaaaatgta 240
aaattagttg gtatttcctt ttgaataagc ttgaccttta acttacttca tagcctatga 300
caactcatat attataaata taaataataa taaatgagtg ctaataataa tttttaatca 360
aaaattctca caagctaacc ttgtaatgat ggtacgccta actgataccc attatgttac 420
catggccagg cataaattcc tgttactggc accgcatcat catcatactg gcattcggtc 480
atattaaaat aatatcacct atctataggg ctgcggctct tttgacacgc acacgtgctg 540
ctcagcagac caagtgaact tatcaatgaa cacccgccac acccacacgg tggcccaagc 600
acctgaccca cccatgccca gtggccagca ccacccgaac tcgcagaagc caatgaatgg 660
ccggcagacc aagaaacacc cactaacatc tcgggagatt cttcgtctgc cgagttcaag 720
aacaacacac aagtggccgt aatgaaatgt gaatagcctt ggcaaacatg agtccggaaa 780
gccaacttat gatggcttct gaatagccaa ccgcatagta cacggctaag agcggggacc 840
aaaaaaggac ctgctctttg ctcgctcgtt ggctttctct ttcgcagcct cggctcggaa 900
aaaggatgcc gacaagcaac aattttcatt gataaaaaca ccagctgctg ggtcctgaaa 960
ctcagtcgta cgcgagtcct cagtcagtca gtttcgagct ctgggatgca acggatgcgg 1020
atgccttaaa tccccgagcc gtttggcggt gttattgaaa attgcgtatt cgccgcgtat 1080
taccaagtaa tagcttgtcc aaagccacta cgctgtcaaa ggacgggctt aaaacgctac 1140

ttgcattttt acgtactata catagataaa ttttccgggt aatcggcgga aattgagtta 1200
aagttttgtg tccggttggc gccagttggc aattggtgct ggccaacttt cggacccgaa 1260
cgatgacaaa ctgaaaagta aactttgtaa agactcagca ccttacaaac aaattcaagt 1320
gtttcgcgaa aaaaaataaa aaattttgaa ttgaagagga cacagttcgg tgaaaaaaaa 1380
atatttgcaa tgcacaatgg aaaatcagca aagtgacctc caagtttacg aaaacagaaa 1440
ctaactcaat aaccatccct gcaggtgagc agaggcaaca tttttttcttt aaaacatttt 1500
ttaccatatt tccgtaaacg gtttgcgtcc tttggcaaat ttatggttgt cttggaataa 1560
aatgaacttt atttacttgg caacaaaacg agtgtttatt cttgtcgcat agggaatcaa 1620
gcagccaaca taaggactat attttgttta attcgattta atgtgagttg cagaattaaa 1680
tttcttttta attcaacacg catctgtaca ccatttgtaa aattacaagc taaagctcag 1740
caaacactaa attttaatca gtgaaacgaa ctgtattctg tcaattttaa tatcaaaagc 1800
atatacattg ctcttaaaaa tattaaactg cgttagtaac gtaaacagaa aatgctatga 1860
aatttgttat aatcaaaatt cacttgaaag agatttcaac aaaattacca atctcgtgtt 1920
ctaaggaatt aatataataa ttataataat atcatagctt gttccctgtc acctaatata 1980
taccatttaa ataaaagtaa tgcctaggta attgtttgtg agcgataagc gcagatttgt 2040
ctgcataatt atgtgttaat ttgattgagg tctactgaag agggccttat ttttataagc 2100
caaaagaaca aattcaaaca aaacacttga gggcggcctt agctggcaat aagcggcttt 2160
tgcctaatgg aaaataagtg gaatgcctag gaattattgg ccaagtggac gggctgctct 2220
ctttctggca atcctgagtc ctttcgactg attcgttttt ggctcctttg gccccatctc 2280
cgcacgcttt gttcggcttg acacacctga gtggcgatgt tccacacagg cactttccca 2340
aacaaatgga acccggcatc tcgcatttgg ttgcccgaac gtaagtccca tgaagtggtc 2400
ttctccagtg gacagccagt cagccgcatt tcttgggggc gctcttctcg agcgcttggc 2460
ctctgagatg gcctctttgg atgtcttgga gatatgcagc tcatcctggc cacccacttc 2520
gccgagccat tccatttcat ttccctggcc aacatacccg tatgaaacga tgagtgctct 2580
tcacagtttt tgcctacatt ccggccactg caccgactga tggttgatgg agacccattc 2640
tccggttagt gttcgtcttg gtcgtcttgg tctgggtctc ggtcctggtc acaatgggaa 2700
ttcgcagctt aaaggcgccg tggctttatt caggccagat tggatacttg acatgttcta 2760
atgcggcgca gatctcagtt tctgcaacgt gctctcccta tctcagttac atgaattagc 2820
attcatcgcc actgcactga gcaaaaaata ggaccacaaa tcgctctaac tgctgtaaat 2880
tgatatccat tcaattattt tatttaagta ttataacaca ttgatagaac attaacatgt 2940
catatatttt tgaacacaat tcagcattac actttaataa agtcgaattt gaatttatcc 3000
aactgacgcg ttgggtgata tcataaattc aatttcgaga atgacaggtc attaaaatgt 3060
tgtcacctac actgtcaaat atttaagtca aaatattaag ttataaaata tcaaaaagct 3120
ttgatagctt tagagtatga tttacgttta aaaggattgt acacgaatac aattttgttt 3180
ttttaagcaa tcagcggcgt gtttctgttt tgccatcata aatcgcagag tttcataatc 3240
cctctaaatt aaaatgtcct ttgggatttt tctgaaattt ttttcaaacg cttcccccctt 3300
tttgtttcct gcatgtttag gccgttaaga gcctgttgat gatggcggca tggtagaact 3360
ttctgttttg tcacagcctt gggctgctgg ctttacattt ggtggaaaaa cgagaccagc 3420
cgagcacgtg cggtccgcgg ggttgggggg ggttgtggaa gctaaggacg tggcaggacg 3480
aaagcgctga gtgcagatga ggactgcaat aaactgccat acataagtga aaattgtgaa 3540
atgaaaactt gcgcgagcat aaatcaatac acccaagggg ggtcatgggc tcccatggca 3600
ggcaacaaaa aaatatatat gtatatatat cagctgggaa atgtgctccc tactcaggtt 3660
tagcgacttg tatttcatgt tcccaaatca gctggtgtta tcctttgatt gccgggcaca 3720
cgacaaaagt ggtggaccag gacacatgtg atgattgctt gcgtgggtcg caggaatgag 3780
ccaactcaga ttggtaatat tgtgtatgaa cgtgatttcg tgactttgca atgtttctat 3840
ttaacaaaaa ctcttgttaa agaaatacac acaaaacatc acttataacc ctttttggat 3900
atctatgagg atttaaggtt tgtaaactaa ggcaatggta taccttaaaa agtctatagt 3960
aaatcgcata tattacatca atttaacgt tcatgctgat cttccagttc caaaatatat 4020
attttgaaat agcttacatt taattaagtg cacttgctat aaaagttttg agcttgcctg 4080
aagtttatta cataaaaaat tcgcagaaag gcggctaaag aatgaaacta atttaagccc 4140
aaatgataaa ctatttgcca agcatgtttc gtttgaatta tgctcccgtt gctctgccta 4200
ccaaaaatgg caaaaactcg caatatgatt attgtggcag cttcttccgg tgaattaaaa 4260
atacaaaaaa ccagctcaaa ttgaaactta attttaataa attgttttca tgctcttttt 4320
cgctgccgtt gtagcgtttt cccgcccatt tcacccgagt ggctggataa ttgtatccgc 4380
gcaatcatcc cgcggcttcc attgccgcag gataaaaacc aaactgcaac gaacagaaca 4440
gaactgagca ctttaagcca tttccactcg taattaaatt ttgagagcgg atgaaattgt 4500
ggcggctaaa taaaaggaac aaaatgcatc tgtgcataca gcaaaaaaaa aacagcaaaa 4560
ataaataaca aaatgcagaa gagcaaaaaa tcgcgcatat ccgctttgta aaaatacgag 4620
tctaagtgat gggttccgtt tttgtttagt agtgctcact agtgaacggc cttgtcgcat 4680
ttggggctga tagtggatgt agtgtggggt gtattgtggt gtgcatgggg gtgtttgaat 4740

```
cgcagaagtc ggcccagccc gttggccgtc aagaggacct ggcctaatcg aatgaatgta 4800
gaatgggcac aaacaagtca cgcctgcctg cgcatttcat caaagaacgg cgaacaagag 4860
gctaacagct cttcgtttct ctctctcttt ctctgctcag ctgggtcctt tgaagatcct 4920
gaatgacact caattggcct aatcggtttc tgaaaagcat agtatgaatt tcatacggta 4980
cgctaatttc gaatataaca aaaagcaaaa acgcttttga tatctaaatt gtatttgcca 5040
ttgaaaagtt tatacctata tgtataatat ggaacctttc taaatatagt aattgatacg 5100
aaattgcaaa aatgatggat gaacatctca acattcaaca ggttgcccag ttacttaatc 5160
aatccatcgg tgttgaaagt gagcacttta tacaaaaata ctcatgagtg gttctaaagc 5220
ggataagagg atttatcttg agctcaaaaa ctgatgcagt gatgcaatat tcaagatgtg 5280
tgccttagta ttcaacattt ttgccagatt tatcgaaacg taaataaata aaaatataca 5340
tgaaatttaa taacattgcg cagcttttg gtgaaagcag ttttaattta ataaagcaat 5400
cgttgcgttg aaactaatat tattttaaat tatacaggat acaaaaaata atatacatat 5460
acatttaaat tttgaattta aaaaggggtg tgacagattt actgaataat aataacttta 5520
aacatgccac ctaacaaaag cgaaaacaat tgtttaatta catttgtgca acttgtggca 5580
tttagtaaaa gcgaatatct aagcaaataa gtcgtgtaat ttgatatgca atatagcaac 5640
caaacaaaaa tgcgtgttat taaaatcatg attactgcca tacgtataga tgtttcatgt 5700
cctgtccctt gttgatttca tgctgataaa ggacatacac ttttaaatct tgctaacttg 5760
gtacttttaa cgagcattcg aaggtatcct gtgcccattt caaaaaagct gtaagctctt 5820
tggcagccaa aaaagctgac aaatttgtgc cggaaaggaa acaacttcct atttgcatga 5880
tttattgcca atggaataac aggttggctt cagttgcgta tgtccctgct attcgatttt 5940
caacgaaatt tcggttggca ttcagtcgga atgaaagtga aaacaaaaag tgctactact 6000
acccttttgt gaaagccccc ggagaaaaga caactttcaa ttggccaatt agtgctattt 6060
aaatggcaag ttacgccgtt ttcaaagggc caactatttt gtttttgaag ttccgcgagt 6120
tgtaaatgta ataagcaaat ttcgcagttc tccataggac acctttacac tgagttcctt 6180
tccagttgct caagttttgc tccttggcag ctgcacttgt tgttcctgtt tgccagaaag 6240
ttagtttctg gccaagaatt tgtctttggc caagtactgc tcacgctcgg tttctttggg 6300
caaaagttgg tatttgagtt tcgcagtgga aaagcacttg cggcagatct tcaaattctg 6360
ccaaacccgg cgggctcacg tggccgcctt ctagttattt aataagtaat taagttccct 6420
aattaaatac attctcgcag agatttacat aataagtaat ctacgccgaa cgtcttcctc 6480
cattgcggcg ttaattaaat ttgcagaccg attgcatttc ggaaacttga tttggtttgc 6540
gtctggctgc cagtcgttaa caaatcatga ttgctgaaat ggtcaacttt gggcatcgca 6600
cttcgtatac acttgagatt aaaaattctt taataagggg aaataacact tatttataga 6660
agtattttta aaatatatac aattaatcta gtcagcgctt ttgacgtaat ttatattgag 6720
ttcgaacaac tcatttgaca acatcaaaga atccaatatt cagcaatttt aaacatactc 6780
acatgcgaag gggaatcga atttccaata atttaaaaag gaatcggaat tttacatttt 6840
ctcgcgcttt tgcggagaag ccgtgatatg gtgactgtat ttatgaaatt tgcttttacg 6900
ccacttccat gaactgaaaa ttggttttca ctgcctgttt ttggtcgata taataatgct 6960
gagatatata tatgatatgt aatgtctttc tatatattgt agcatttacg aagggttata 7020
gcttgatata aaaaaagaaa gaaactttga acaaactcaa gtatataaaa aggtttcaat 7080
aaaatatata aagtccatta gctgtagctt acaggagcat aagttttctt gtttaagacg 7140
aagctttatc gaactttatc ttatagatac atatatctat tactaggata agtaagtaat 7200
tacaattatt tatgcacgct gcattgacat atgccaaaca cttaaagtat tcccacccag 7260
ccgcacttgc tacatttatc agcaattgaa ttaatttcag gttgcggttc agctgagttt 7320
ggtttgtaaa atacattata atgtttaata atcagcagag agcaggaaaa atatgtatta 7380
agttcggcat tgtgttaaac tttaagcgaa tagttaacac agtgttattg gttacccgct 7440
gcttttaata actttccctt tcgatgccca cacaacctga gttattttcc ttaatggtct 7500
gaaaaaatga agaaaaaaaa gtttggctca agtttattgg tttcgcgctg ctttcatcaa 7560
ttgctgcacg caaatgcttc tcaaaaaaaa aaaaaaaaaa actaactctt catttatggc 7620
atttcttagt tgcgaataaa ttccgattta agctggttga aaaaaaccca tcctttcaac 7680
ggtgactaaa catattttaa gcagaccgct caaactcaaa ccgtgggtgg ttaattaaat 7740
aagagcatcg attaagtcaa ctgtcttttt atctgtccca cttttagggt cacccggtgg 7800
acagctaatt gcagctaacc ccattggcat ttagcaaacg gaaaatgcat cgtgaccgct 7860
gcaaccacat tgcaccacta ccaccagcag catcagcgcc accacttcca cctccagcag 7920
cagcaaggga atccgcaccg ctgaaactga atctgaaggc atcccagtgg cggcgacaca 7980
cggaatcctt gccaaatgcg gccaaatgcg aggcgctggc aagtaaataa gcgtaatcct 8040
ttgcgggcac gtaccaccac tactacctcc tgttcggagt agctcgtaaa tatttggcca 8100
agtcagtcat ggaggccatc atcatgacga ccctgcccaa cctgacaacg gatgcaggtg 8160
acagcagctt ctggctaacc ggtgccctct cgctctccga gatgttggcc aactcgagcc 8220
acagccattc cactgggagc acaacctcga cggctgggag ttcggccacc gaatcatccg 8280
ctgtgaatgt aggcaaggac cacgacaagc acgtgaatga cagcgtttcc acgggtctga 8340
```

ggtaagtgca tgtgcctagg cttttccatt tccatggcaa cacacacaca gattccatta 8400
agctcatgac tgcgggctca acaagtcatc gcgataagct cagcttcatt ttccaggaga 8460
tgcattcaac taatgtgctt agaattctac tattgtacat aactaacata tagggaatat 8520
aatcagatta cttagaacct cagaacagtc aatacaatcc acttaatata aatgaattaa 8580
aagtatctat cgaaactgaa aagaagagga tcattagctt gggtacaaca tttataagaa 8640
ttttggctta tgaagttgca aaaagaattc ttaaaaaatt caatacaatg tcaaatagtt 8700
ctgaaaccat aaattattaa caattaaata ttaaattatt tatattgatt taattcatac 8760
atttaaacat aagttacact actgcaagga ggaaagtatc agatgacttg agctatactt 8820
tcataaactt agcttagagc aatctaaacc aaattgaaaa ctcattgaaa cccattttaa 8880
ttgaataaat tgaaacattt acttttccca tttccccagc aattacagta attaccccag 8940
ctacattcac tacagggaca agtacgacct gagctacatt gccaaggtga atccctttg 9000
gctacagttc gagccaccta agtcgagcac cttcctgatt atggccgccc tgtattgcct 9060
gatttcggtg gtgggatgtg tgggcaatgc gttcgtcatc tttatgtttg ccaaccgcaa 9120
atctctgcgg actcctgcca atatcctggt gatgaatctg gccatctgcg actttctgat 9180
gctcatcaaa tgcccgattg ccatttacaa taacatcaaa gagggtccag ctttaggaga 9240
cattggtaag cttaaaactt ggggaaataa cttactaaat atctatataa atgaaattca 9300
ttcggcttgc atttatgtca aaaatcgatc aacaaaagca atctacataa taaataaaat 9360
taaattggcc caaaatgctt ggactgattt cagctgtttc ttgcagaaat tcatttgtgg 9420
gattctcttt tcccgaaagc aattaataga atttaaatga aacaataaaa gttcaatttg 9480
tagcctacac atcttggttc atgaaaaaca tattccaaaa tgtaatggct attaaaattt 9540
caccgagcct tttgtttgcc ttgaaacggg tggttcatta aaaatcctac tgactttccg 9600
tcccacagcc tgtcgcctct atggatttgt gggtggccta agtggcacct gtgccatcgg 9660
taccctcacc gccatcgctt tggatcggta caatgtggtg gtgcatccac tgcaaccgct 9720
gagacgctgc tcccgcctgc gatcctacct gatcatcctg ctgatctggt gctacagctt 9780
cctgttcgct gtgatgccgg ccctggatat cggactatct gtctatgtgc cagagggctt 9840
cctcaccacc tgcagcttcg attatctgaa caaggagatg ccagctcgca ttttcatggc 9900
actgttcttt gtggctgcct actgcattcc actgacctcc attgtgtact cctacttcta 9960
tatactgaag gtggtcttca cagcgagtcg aatacaatcg aacaaggata aggccaagac 10020
ggagcagaag ttggccttca ttgtggcggc tatcattggt ttgtggttcc tggcctggtc 10080
accgtacgcc atcgttgcca tgatgggtgt cttcggcctg gagaggcaca taacgccact 10140
gggatcgatg atacccgcgc tcttctgcaa gacagctgcc tgcgtggatc cctacttgta 10200
tgcggccacc catccgcggt ttcgtgtcga ggtgcgaatg ctcttctacg gacgcggagt 10260
gctgcgaagg gtctccacga ccagatcatc gtacatgacc cgatcacgat cgtcgttcac 10320
ccaccgattg aggacgagca ccaccggtga gggcggaatg ggtgatcacc gcatggaaaa 10380
ctatctgatg aacaacaatc tgatgatggt gcccgaggag acggaggaga acgaggagat 10440
cgttgtggtg gccgagatca acaattcgat tagcagcgtc atggaacaga gtaagttctg 10500
aaggctgacg attatgccag ttaagagaac ggaaattctg tcaatctgtt tcgaaccaac 10560
taacataaac atttttaatc agttttaagt tttacagacg tgttttatat aatttgaaaa 10620
ttagtcaaac ctatgatcac gatgatactg aaaactttgc ttggaatatc tgaaactcgc 10680
ctgtttaatt tccacttttt ccaattgaaa atttcaattt gatttttgcc caggtaattt 10740
ataaaaaaac ctttcctatc tcgaaaagta aactgttcgc tgttaattga aaatgtttgt 10800
gcactgcact gcaaggcata tcaaaatgga agctatcaat aatcgtaaat tgaatggatc 10860
ggactcaagt ttcattgttg aactaatcta atctaatttg tatgtagaaa taatagaccg 10920
aactgaaatc agcaatggca gcaaagtaat ttcgctcggc aatgatatat gccaatggca 10980
ttgttcatct gagtgaatgc gttgtcattg atggacctac aattgaaatg aaaataataa 11040
taataataaa tagaaatgaa gcagggaaat ccgccgagca gtgggaaatt ggcaaaggca 11100
caacgacgac acagtggtga tagtaaacat atttaatgag cttatgtgat gtttagacag 11160
ggggagtagg cgatcagagc agggaatgga ttagttagcc tatatacata tatatataca 11220
caggcatatg tacatatgta taagacatcg gaatagcagg ccaattagag aggcagggcc 11280
gaaattagta caaagaaaga gggtggcgtg ggtaggggcc gaagtggaaa ttcaatagtc 11340
agattataca tcatttcagg gaaagtcact cggacccaag gtcaagccgt ggtattggat 11400
gttaacaagc ggatgtgcag attgtttgca ttgtagttta atattggaat aactcacaat 11460
caaagcagcc tgcggacatc ttgtatttct gaaatggagt c 11501

<210> 35
<211> 2177
<212> DNA
<213> Drosophila

<400> 35
ctgggatgca acggatgcgg atgccttaaa tccccgagcc gtttggcggt gttattgaaa 60
attgcgtatt cgccgcgtat taccaagtaa tagcttgtcc aaagccacta cgctgtcaaa 120
ggacgggctt aaaacgctac ttgcattttt acgtactata catagataaa ttttccgggt 180
aatcggcgga aattgagtta aagttttgtg tccggttggc gccagttggc aattggtgct 240
ggccaacttt cggacccgaa cgatgacaaa ctgaaaagta aactttgtaa agactcagca 300
ccttacaaac aaattcaagt gtttcgcgaa aaaaaataaa aaattttgaa ttgaagagga 360
cacagttcgg tgaaaaaaaa atatttgcaa tgcacaatgg aaaatcagca aagtgacctc 420
caagtttacg aaaacagaaa ctaactcaat aaccatccct gcagggtcac ccggtggaca 480
gctaattgca gctaacccca ttggcattta gcaaacggaa aatgcatcgt gaccgctgca 540
accacattgc accactacca ccagcagcat cagcgccacc acttccacct ccagcagcag 600
caagggaatc cgcaccgctg aaactgaatc tgaaggcatc ccagtggcgg cgacacacgg 660
aatccttgcc aaatgcggcc aaatgcgagg cgctggcaac tcgtaaatat ttggccaagt 720
cagtcatgga ggccatcatc atgacgaccc tgcccaacct gacaacggat gcaggtgaca 780
gcagcttctg gctaaccggt gccctctcgc tctccgagat gttggccaac tcgagccaca 840
gccattccac tgggagcaca acctcgacgg ctgggagttc ggccaccgaa tcatccgctg 900
tgaatgtagg caaggaccac gacaagcacg tgaatgacag cgtttccacg ggtctgagca 960
attacagtaa ttaccccagc tacattcact acagggacaa gtacgacctg agctacattg 1020
ccaaggtgaa tccctttgg ctacagttcg agccacctaa gtcgagcacc ttcctgatta 1080
tggccgccct gtattgcctg atttcggtgg tgggatgtgt gggcaatgcg ttcgtcatct 1140
ttatgtttgc caaccgcaaa tctctgcgga ctcctgccaa tatcctggtg atgaatctgg 1200
ccatctgcga ctttctgatg ctcatcaaat gcccgattgc catttacaat aacatcaaag 1260
agggtccagc tttaggagac attgcctgtc gcctctatgg atttgtgggt ggcctaagtg 1320
gcacctgtgc catcggtacc ctcaccgcca tcgctttgga tcggtacaat gtggtggtgc 1380
atccactgca accgctgaga cgctgctccc gcctgcgatc ctacctgatc atcctgctga 1440
tctggtgcta cagcttcctg ttcgctgtga tgccggccct ggatatcgga ctatctgtct 1500
atgtgccaga gggcttcctc accacctgca gcttcgatta tctgaacaag gagatgccag 1560
ctcgcatttt catggcactg ttctttgtgg ctgcctactg cattccactg acctccattg 1620
tgtactccta cttctatata ctgaaggtgg tcttcacagc gagtcgaata caatcgaaca 1680
aggataaggc caagacggag cagaagttgg ccttcattgt ggcggctatc attggtttgt 1740
ggttcctggc ctggtcaccg tacgccatcg ttgccatgat gggtgtcttc ggcctggaga 1800
ggcacataac gccactggga tcgatgatac ccgcgctctt ctgcaagaca gctgcctgcg 1860
tggatcccta cttgtatgcg gccacccatc cgcggtttcg tgtcgaggtg cgaatgctct 1920
tctacggacg cggagtgctg cgaagggtct ccacgaccag atcatcgtac atgacccgat 1980
cacgatcgtc gttcacccac cgattgagga cgagcaccac cggtgagggc ggaatgggtg 2040
atcaccgcat ggaaaactat ctgatgaaca acaatctgat gatggtgccc gaggagacgg 2100
aggagaacga ggagatcgtt gtggtggccg agatcaacaa ttcgattagc agcgtcatgg 2160
aacagagtaa gttctga 2177

<210> 36
<211> 551
<212> PRT
<213> Drosophila

<400> 36
Met His Arg Asp Arg Cys Asn His Ile Ala Pro Leu Pro Pro Ala Ala
1 5 10 15
Ser Ala Pro Pro Leu Pro Pro Pro Ala Ala Ala Arg Glu Ser Ala Pro
20 25 30
Leu Lys Leu Asn Leu Lys Ala Ser Gln Trp Arg Arg His Thr Glu Ser
35 40 45
Leu Pro Asn Ala Ala Lys Cys Glu Ala Leu Ala Thr Arg Lys Tyr Leu
50 55 60
Ala Lys Ser Val Met Glu Ala Ile Ile Met Thr Thr Leu Pro Asn Leu
65 70 75 80
Thr Thr Asp Ala Gly Asp Ser Ser Phe Trp Leu Thr Gly Ala Leu Ser
85 90 95
Leu Ser Glu Met Leu Ala Asn Ser Ser His Ser His Ser Thr Gly Ser
100 105 110

```
Thr Thr Ser Thr Ala Gly Ser Ser Ala Thr Glu Ser Ser Ala Val Asn
        115                 120                 125
Val Gly Lys Asp His Asp Lys His Val Asn Asp Ser Val Ser Thr Gly
    130                 135                 140
Leu Ser Asn Tyr Ser Asn Tyr Pro Ser Tyr Ile His Tyr Arg Asp Lys
145                 150                 155                 160
Tyr Asp Leu Ser Tyr Ile Ala Lys Val Asn Pro Phe Trp Leu Gln Phe
                165                 170                 175
Glu Pro Pro Lys Ser Ser Thr Phe Leu Ile Met Ala Ala Leu Tyr Cys
            180                 185                 190
Leu Ile Ser Val Val Gly Cys Val Gly Asn Ala Phe Val Ile Phe Met
        195                 200                 205
Phe Ala Asn Arg Lys Ser Leu Arg Thr Pro Ala Asn Ile Leu Val Met
    210                 215                 220
Asn Leu Ala Ile Cys Asp Phe Leu Met Leu Ile Lys Cys Pro Ile Ala
225                 230                 235                 240
Ile Tyr Asn Asn Ile Lys Glu Gly Pro Ala Leu Gly Asp Ile Ala Cys
                245                 250                 255
Arg Leu Tyr Gly Phe Val Gly Gly Leu Ser Gly Thr Cys Ala Ile Gly
            260                 265                 270
Thr Leu Thr Ala Ile Ala Leu Asp Arg Tyr Asn Val Val Val His Pro
        275                 280                 285
Leu Gln Pro Leu Arg Arg Cys Ser Arg Leu Arg Ser Tyr Leu Ile Ile
    290                 295                 300
Leu Leu Ile Trp Cys Tyr Ser Phe Leu Phe Ala Val Met Pro Ala Leu
305                 310                 315                 320
Asp Ile Gly Leu Ser Val Tyr Val Pro Glu Gly Phe Leu Thr Thr Cys
                325                 330                 335
Ser Phe Asp Tyr Leu Asn Lys Glu Met Pro Ala Arg Ile Phe Met Ala
            340                 345                 350
Leu Phe Phe Val Ala Ala Tyr Cys Ile Pro Leu Thr Ser Ile Val Tyr
        355                 360                 365
Ser Tyr Phe Tyr Ile Leu Lys Val Val Phe Thr Ala Ser Arg Ile Gln
    370                 375                 380
Ser Asn Lys Asp Lys Ala Lys Thr Glu Gln Lys Leu Ala Phe Ile Val
385                 390                 395                 400
Ala Ala Ile Ile Gly Leu Trp Phe Leu Ala Trp Ser Pro Tyr Ala Ile
                405                 410                 415
Val Ala Met Met Gly Val Phe Gly Leu Glu Arg His Ile Thr Pro Leu
            420                 425                 430
Gly Ser Met Ile Pro Ala Leu Phe Cys Lys Thr Ala Ala Cys Val Asp
        435                 440                 445
Pro Tyr Leu Tyr Ala Ala Thr His Pro Arg Phe Arg Val Glu Val Arg
    450                 455                 460
Met Leu Phe Tyr Gly Arg Gly Val Leu Arg Arg Val Ser Thr Thr Arg
465                 470                 475                 480
Ser Ser Tyr Met Thr Arg Ser Arg Ser Ser Phe Thr His Arg Leu Arg
                485                 490                 495
Thr Ser Thr Thr Gly Glu Gly Gly Met Gly Asp His Arg Met Glu Asn
            500                 505                 510
Tyr Leu Met Asn Asn Asn Leu Met Met Val Pro Glu Glu Thr Glu Glu
        515                 520                 525
Asn Glu Glu Ile Val Val Val Ala Glu Ile Asn Asn Ser Ile Ser Ser
    530                 535                 540
Val Met Glu Gln Ser Lys Phe
545                 550
```

<210> 37
<211> 6472

<212> DNA
<213> Drosophila

<400> 37
tatgcgcaac tcgacaaaca tcttcctaac caacctgagc attgcggacc tgctggtgct 60
gctcgtctgc acgcccaccg tcctcgtgga ggtcaacact cgcccggaga cctgggtact 120
tggccacgag atgtgtgagt ggcttcagct ccagtgctcc tggtctattt gttctcccca 180
acctaattaa ctgctgttct atcccaggca aggctgtgcc gttcgtggag ctcacggtgg 240
cccacgccag tgtgctgacc atcttggcca tctcgttcga gcgctattac gcaatctgcg 300
agccattaaa ggccggctac gtctgcacca aggggcgggc catcctcatc tgcgtcctgg 360
cctggggcat tgctgcgctc tttacgaggt aagcgattag gctcgtactt gatttcttta 420
aatttatttg taaactagca ttatacttta aatatcttaa ttagtttggt cgcttttcat 480
cgcatgcttt ttggtgtttg tagtcctaca tttcgtatgt acatatggta ttatcctgtt 540
tagatctact tcgtttgtgg cctatatccg tttattcatt tcgcccagtc aagcattaag 600
ttcctcgctc attaattata agtgtcaatc gccgtcattt taatgatttg tgttttgtta 660
ttgggttttg cacatttttc cccgttttgt ttattttcca ttatgacgtg gaatttctgg 720
tcttggtgtt tggccttttg tgcgccatct acgctggcgg cgtgtccatt tgtccattta 780
catgctaaac tgcgactgtc gcctatctat tccatttcca ctcctgtcct accacttctc 840
ctgcttcact cgcttcatcg ttgccattgc catttccatt cccaatccca ttcccattga 900
catccgtcgc ctcaagttgc ggacggcttt tgtccaagtt ccggaagctt gccgttatca 960
gcgccctgtc atctttaata agcaccgatg tccgctatca tggatttcat ttcattgaca 1020
ctttattatg gtgcgctccg tccaaggacg agcggcgacc ctttttaatt ttgcaaatct 1080
tgccgcgtcc ataacaatga aataaacagc acattgtcga gctggccacc gacccctcag 1140
cactccggga gtcatggggt catctatggt tggcagctgg gtccgcagga cacccgccta 1200
cgggccgaga cccattaatt acgccacaca agaagcccga ggccaaaaaa aaaaaaaaaa 1260
aaacgggcaa acatttttaa acacattcaa acagacgcgc cacacgacgc ccatgggcgc 1320
aaaaggaggc gtgtcaaata gctgacgcac gtaccaaaac gcccggatca tctttagtgg 1380
gaaaaaaaac acggccatcg cctgttgact gtcatggcat ttcgttttcg aggtctagcc 1440
agggaggatt tgggctgggg ctgtgtttaa tcagcggctc ggataactca aaataatcaa 1500
atatttaaac ttgtctgttt gctggatata tttacaaagt caacagagta tatcccacca 1560
taagttcgcc atctactcac ttatcccctt gtaagtgaaa tgtaatcaca tcaatgaaat 1620
ttggaatatt ttgagcttta gttaaatgta ttttaagatg tatcgatata actcatacag 1680
atactttttt aaaatgagtg gtttcatttt tattaaattt ctgttattat aataatttta 1740
aatgcaatgt tgatattata tttcaaattg gcttaacgcc tacagattag aatacactaa 1800
agatacacgg tatacatttt ttttctagac agaaatatac tttaattcaa ctaaattaat 1860
atatatatgt acctctccca tcttttagct attaaatagt cgaggaactc gactatagta 1920
ttatatccgg ttttggttta aaaaccaatc attgtttttt tttaacgaca cagtactatt 1980
taatatgggc ttcaacatat attttaaatg atgaacccaa gaaatgaaaa aaccataaaa 2040
acaagcggtc atgaaaatcg cgtcctggcc agactggtac attggaaact tgaacttcat 2100
cccagactgg gttaaccttt tgcccatcgc cattctctca tcatcacaaa cagcttctct 2160
tcctgcctgc tgaaacctct tctttttccc gccactcttg gccttttcct ccaaaatcca 2220
accgtatcca ttatccatta tccacgcccc cggtaacccc gccacctgat tacgtttctt 2280
gtttgctttc tctccccaat ctgcaatatc cccgagcag ccccattctg tgggtggccg 2340
agtacaagct ggccgagtac attgatggat cgtcggtggc cgtgtgcctg acccaggcca 2400
tcagcgactg gacgctggcc ttcttcctga tgaccatctc ggtgttcttc gtggtgccgt 2460
tcgtgaccct ggtggtgctg tacggcatca tcgcccggaa tctggtctcc aacagggcgg 2520
ccatgctgcg cgcccgtccc acgaagccgg agcttagtct gaaggcccgc aagcaggtgg 2580
tcctaatgct cggcgccgtt gtgctgtcct tcttcgtctg cctgctgccc ttccgcgtcc 2640
tcaccctgtg gatcatcctc agcacggacc aaacactgca cgatctggga ttggtgcgct 2700
actacagctt gctgtacttc tgtaggatca tgttgtacct caactcggcc atgaatccga 2760
ttctttacaa cctgatgtcg accaagtttc gaaggggttt caagcggctt tgtcaggacg 2820
cagggcgatt actgctggaa ttggtgacat tgggaagaag gaaggaagac tcttctcgcg 2880
ggcgcagagg taccttgtca ctgggcatgg gcaccaatac gaacacgaat accaactcct 2940
caaatgccac aggagccacg agctccagca ttctctcgag aagctccaat cgcaggtgca 3000
gcgaggatat cagtcgcacc cgccttaaga tcgagatgca gatgccatgt ggcagcgacc 3060
tggaggccat ggccatgctg cagcattcca ccttgggcaa gggaattgcc aggcgagtga 3120
gtgatagccg tctcatgcct ctgagaaatc atcaaccacg acgccacaag ccgcaaataa 3180
gtttcgacga ggagtcactg gaggagaata agcgtagtga agcgaaaatc ccaacaaaat 3240
gtcgagaaaa gcttcctgga atagcaagag aaatagtaaa tctgaccgaa aataccctct 3300
aagcgagatg ggcgtgaaaa atcgccttta acccgtttac acactcgctt cacttggctg 3360

```
gaatgtgccg atgacgctga tcccgatccc gattcctatt cccgctcctg gtaaacacaa 3420
atcgcttcct gaacgctttt ggcagccata atcgatttca gtcacccgct gacagctgcc 3480
gagaatcccg aatcgtagtt gggctgccgc cttggacatt tccgtaccaa aaacaatcaa 3540
tttttcgggc tattaaactc gaatgaatgt gccaaagata aataacagca acgacagccg 3600
aacagcacta aaactaccac tactggtcag ggttaggagt ccagatatcg actggcagta 3660
aggggtaagc cattaggtgg ccttcgagga gcggagctgg caaatggaaa tggcttaata 3720
tgcgtttggg cgaaccatca ggtcggcgga taaacatgat ttaggtgtat tgaaatagtt 3780
tccgtacgcc ggaccgtaca aaatggccgc cattacacgg gcggccatca ccggaagtcg 3840
aaaaaactgt tgcctacttg ccggcagctc aaataaaaag cttagctcaa tatgcacaca 3900
atgtccacat gtaaacacag ctgtacatac acatacatac ccccacacgc acacatagcc 3960
aaatattaat atttatattt gtacgaacat atttatatac acatatttac atatatatag 4020
tctgcttata aatatgtaca aggaaccaca cgcacacact cacacacaca cgagagtttg 4080
taagtaaata aattttttat gacaactcat ggcgggactt ttccacatta ccccatccaa 4140
tcggaatctg taatttatat aattctaagg aaatcggaca aatgttcgtc ctctgtaaat 4200
atacgtattc gtaccatgcg cgtagctttt taatttgttt tcgtagccct atgacgtaga 4260
ctcccacttt tgtagagtaa aaagtaataa gtagcttgaa cccacaccat cctgcgagta 4320
tgatgagtat gagcagacgc tgccgtttcc gtttccgctg tgtgattaag ttgagtgctt 4380
cgccccccgaa acttttccag tagttagccg ttgagtagct tttaaatatt tgtttgtttc 4440
ctgcccgcag cgcgataaaa ctttgcagac agcgctagtc actctagtcc ccgctaaaac 4500
tgataatatg attcttgcag tcccatcatc gccatatcca cctacagcgt ggaaccctac 4560
ggggacggaa ccgatgctcc cgtttgcacc accgccgccg atggtttctg gtcgatcttc 4620
tacttcgtgg gctgcatcac ggtgtttttc ttcctgccct tcggcatcct ggttcttcta 4680
tacgcggcca tcgcttacaa gctgctccgt cccaacaacg ccttccaccg accaacctcc 4740
ccgcagccac agcagccgtc cggcggagcc accagtggct cctcacaggt gcccagcacc 4800
aagggtaaca gccatcaaca aagcaacggg atgaggaagc atcgcaagca ggtaatcttc 4860
atgctggtgg ccgtggtgtc tagctttttt gtttgccttc tacccttccg ggccttcacc 4920
ctctgggtga tcctggccag cgccgaggat gtcgagggtc tgggcattgc cggctactac 4980
aacctgctgt acttctcgcg cttcatgctc tacctaaact cagccatgaa cccgatcctc 5040
tacaacctga tgtcctccaa atttcgcagc ggcttctggc ggctgctgct cacttgtctg 5100
ggccaacggc cacatcacca ccatcgccac cactaccatc agaggcagca tccaacggca 5160
ggcggaagtg ggcgcaatgc gtccacgcga caggaacagg atgccgagga aggagctgcg 5220
ctggcgggaa cgaccagcgc ccgacatcca cgtcgcacac tccgccgcga ggccaccttc 5280
ttgatcaact ccatatccac ctcctcgggt acggatcgca ccacatcatc atcggcgtgg 5340
cgcagcaaca gtctgtccat ttccggtctg agcgaacggg agcgcggcat actgggagcc 5400
gctatcatcg gcacaacggc tgccaccgtt acaaccgcct gtctgcagga gcgacgcgcc 5460
agcaagatct gaatcccatc ctaaatatgc atccttcgag tccctgtcca ctgccagcaa 5520
ccgttaaagt atttagttga attccaacag caaatagtca tagttttaga aaaatttttgt 5580
tgtaaagact gaagtacctc agaagaagtc aataaacgca agagaagatg tttttgaaaa 5640
tcgaaactaa ttttcttcgg gcttcgaaaa attgaagttt gcaagattcg aaatagcttt 5700
aacttgattt ttagaacaga aacgagtatg aacttttcat acgatcattt ttatcgcacc 5760
tattgtgatt taatgggaga taggaggaga ccaattggga tatgcaaaaa taacgcaact 5820
attattatct ttgtatgtac gattaatttt aatctctaaa gagtcaaagt acttcaataa 5880
ttttacaaat ctgtcggagg gaatgtaagc ttgcccaaat caaatggaag ctacaatcgt 5940
tgtcagatgg agatatttta tgttagaagg aatatctcgg atacgaaata tttaatttaa 6000
aattcctttt catattattc tttttaaaaa ataataagtc ctgtatgtat acaatgcaat 6060
acataaaggt agagtatgtc ctaaaaagaa aaatatcttt ccgaatatat gtatataatt 6120
tataaaataa gtggatacat acatttttcg taagtatcct agaagatttg ccaacgtatc 6180
aggtttacaa aaattgatta cctatgtgca gttaacttaa cggcagtgca ctaactaatc 6240
tgatttgtga aataatgcag atataaaaca tttttaaata gaaatatatt tatacatctg 6300
attcttcaac atctaagcac tcgtatagag tgctttgaat atattctaaa tacattggca 6360
tgatccaagt tttaaatcag ccaaaaactt gctccaaaaa caggaagatt ctacacaaat 6420
cgcctcgccc ttaaagtttc tgccaagtat ttgggtgcgt gcggatcagt tc         6472
```

<210> 38
<211> 2196
<212> DNA
<213> Drosophila

<400> 38
atgcgcaact cgacaaacat cttcctaacc aacctgagca ttgcggacct gctggtgctg 60
ctcgtctgca cgcccaccgt cctcgtggag gtcaacactc gcccggagac ctgggtactt 120
ggccacgaga tgtgcaaggc tgtgccgttc gtggagctca cggtggccca cgccagtgtg 180
ctgaccatct tggccatctc gttcgagcgc tattacgcaa tctgcgagcc attaaaggcc 240
ggctacgtct gcaccaaggg gcgggccatc ctcatctgcg tcctggcctg ggcattgct 300
gcgctcttta cgagccccat tctgtgggtg gccgagtaca agctggccga gtacattgat 360
ggatcgtcgg tggccgtgtg cctgacccag gccatcagcg actggacgct ggccttcttc 420
ctgatgacca tctcggtgtt cttcgtggtg ccgttcgtga ccctggtggt gctgtacggc 480
atcatcgccc ggaatctggt ctccaacagg gcggccatgc tgcgcgcccg tcccacgaag 540
ccggagctta gtctgaaggc ccgcaagcag gtggtcctaa tgctcggcgc cgttgtgctg 600
tccttcttcg tctgcctgct gcccttccgc gtcctcaccc tgtggatcat cctcagcacg 660
gaccaaacac tgcacgatct gggattggtg cgctactaca gcttgctgta cttctgtagg 720
atcatgttgt acctcaactc ggccatgaat ccgattcttt acaacctgat gtcgaccaag 780
tttcgaaggg gtttcaagcg gctttgtcag gacgcagggc gattactgct ggaattggtg 840
acattgggaa gaaggaagga agactcttct cgcgggcgca gaggtacctt gtcactgggc 900
atgggcacca atacgaacac gaataccaac tcctcaaatg ccacaggagc cacgagctcc 960
agcattctct cgagaagctc caatcgcagg tgcagcgagg atatcagtcg cacccgcctt 1020
aagatcgaga tgcagatgcc atgtggcagc gacctggagg ccatggccat gctgcagcat 1080
tccaccttgg gcaagggaat tgccaggcga gtgagtgata gccgtctcat gcctctgaga 1140
aatcatcaac cacgacgcca caagccgcaa ataagtttcg acgaggagtc actggaggag 1200
aataagcgta ggttaggagt ccagatatcg actggcagta agggtcccat catcgccata 1260
tccacctaca gcgtggaacc ctacggggac ggaaccgatg ctcccgtttg caccaccgcc 1320
gccgatggtt tctggtcgat cttctacttc gtgggctgca tcacggtgtt tttcttcctg 1380
cccttcggca tcctggttct tctatacgcg gccatcgctt acaagctgct ccgtcccaac 1440
aacgccttcc accgaccaac ctccccgcag ccacagcagc cgtccggcgg agccaccagt 1500
ggctcctcac aggtgcccag caccaagggt aacagccatc aacaaagcaa cgggatgagg 1560
aagcatcgca agcaggtaat cttcatgctg gtggccgtgg tgtctagctt ttttgtttgc 1620
cttctaccct tccgggcctt caccctctgg gtgatcctgg ccagcgccga ggatgtcgag 1680
ggtctgggca ttgccggcta ctacaacctg ctgtacttct cgcgcttcat gctctaccta 1740
aactcagcca tgaacccgat cctctacaac ctgatgtcct ccaaatttcg cagcggcttc 1800
tggcggctgc tgctcacttg tctgggccaa cggccacatc accaccatcg ccaccactac 1860
catcagaggc agcatccaac ggcaggcgga agtgggcgca atgcgtccac gcgacaggaa 1920
caggatgccg aggaaggagc tgcgctggcg ggaacgacca gcgcccgaca tccacgtcgc 1980
acactccgcc gcgaggccac cttcttgatc aactccatat ccacctcctc gggtacggat 2040
cgcaccacat catcatcggc gtggcgcagc aacagtctgt ccatttccgg tctgagcgaa 2100
cgggagcgcg gcatactggg agccgctatc atcggcacaa cggctgccac cgttacaacc 2160
gcctgtctgc aggagcgacg cgccagcaag atctga 2196

<210> 39
<211> 731
<212> PRT
<213> Drosophila

<400> 39
Met Arg Asn Ser Thr Asn Ile Phe Leu Thr Asn Leu Ser Ile Ala Asp
1 5 10 15
Leu Leu Val Leu Leu Val Cys Thr Pro Thr Val Leu Val Glu Val Asn
20 25 30
Thr Arg Pro Glu Thr Trp Val Leu Gly His Glu Met Cys Lys Ala Val
35 40 45
Pro Phe Val Glu Leu Thr Val Ala His Ala Ser Val Leu Thr Ile Leu
50 55 60
Ala Ile Ser Phe Glu Arg Tyr Tyr Ala Ile Cys Glu Pro Leu Lys Ala
65 70 75 80
Gly Tyr Val Cys Thr Lys Gly Arg Ala Ile Leu Ile Cys Val Leu Ala
85 90 95
Trp Gly Ile Ala Ala Leu Phe Thr Ser Pro Ile Leu Trp Val Ala Glu
100 105 110
Tyr Lys Leu Ala Glu Tyr Ile Asp Gly Ser Ser Val Ala Val Cys Leu

115 120 125
Thr Gln Ala Ile Ser Asp Trp Thr Leu Ala Phe Phe Leu Met Thr Ile
130 135 140
Ser Val Phe Phe Val Val Pro Phe Val Thr Leu Val Val Leu Tyr Gly
145 150 155 160
Ile Ile Ala Arg Asn Leu Val Ser Asn Arg Ala Ala Met Leu Arg Ala
165 170 175
Arg Pro Thr Lys Pro Glu Leu Ser Leu Lys Ala Arg Lys Gln Val Val
180 185 190
Leu Met Leu Gly Ala Val Val Leu Ser Phe Phe Val Cys Leu Leu Pro
195 200 205
Phe Arg Val Leu Thr Leu Trp Ile Ile Leu Ser Thr Asp Gln Thr Leu
210 215 220
His Asp Leu Gly Leu Val Arg Tyr Tyr Ser Leu Leu Tyr Phe Cys Arg
225 230 235 240
Ile Met Leu Tyr Leu Asn Ser Ala Met Asn Pro Ile Leu Tyr Asn Leu
245 250 255
Met Ser Thr Lys Phe Arg Arg Gly Phe Lys Arg Leu Cys Gln Asp Ala
260 265 270
Gly Arg Leu Leu Leu Glu Leu Val Thr Leu Gly Arg Arg Lys Glu Asp
275 280 285
Ser Ser Arg Gly Arg Arg Gly Thr Leu Ser Leu Gly Met Gly Thr Asn
290 295 300
Thr Asn Thr Asn Thr Asn Ser Ser Asn Ala Thr Gly Ala Thr Ser Ser
305 310 315 320
Ser Ile Leu Ser Arg Ser Ser Asn Arg Arg Cys Ser Glu Asp Ile Ser
325 330 335
Arg Thr Arg Leu Lys Ile Glu Met Gln Met Pro Cys Gly Ser Asp Leu
340 345 350
Glu Ala Met Ala Met Leu Gln His Ser Thr Leu Gly Lys Gly Ile Ala
355 360 365
Arg Arg Val Ser Asp Ser Arg Leu Met Pro Leu Arg Asn His Gln Pro
370 375 380
Arg Arg His Lys Pro Gln Ile Ser Phe Asp Glu Glu Ser Leu Glu Glu
385 390 395 400
Asn Lys Arg Arg Leu Gly Val Gln Ile Ser Thr Gly Ser Lys Gly Pro
405 410 415
Ile Ile Ala Ile Ser Thr Tyr Ser Val Glu Pro Tyr Gly Asp Gly Thr
420 425 430
Asp Ala Pro Val Cys Thr Thr Ala Ala Asp Gly Phe Trp Ser Ile Phe
435 440 445
Tyr Phe Val Gly Cys Ile Thr Val Phe Phe Phe Leu Pro Phe Gly Ile
450 455 460
Leu Val Leu Leu Tyr Ala Ala Ile Ala Tyr Lys Leu Leu Arg Pro Asn
465 470 475 480
Asn Ala Phe His Arg Pro Thr Ser Pro Gln Pro Gln Gln Pro Ser Gly
485 490 495
Gly Ala Thr Ser Gly Ser Ser Gln Val Pro Ser Thr Lys Gly Asn Ser
500 505 510
His Gln Gln Ser Asn Gly Met Arg Lys His Arg Lys Gln Val Ile Phe
515 520 525
Met Leu Val Ala Val Val Ser Ser Phe Phe Val Cys Leu Leu Pro Phe
530 535 540
Arg Ala Phe Thr Leu Trp Val Ile Leu Ala Ser Ala Glu Asp Val Glu
545 550 555 560
Gly Leu Gly Ile Ala Gly Tyr Tyr Asn Leu Leu Tyr Phe Ser Arg Phe
565 570 575
Met Leu Tyr Leu Asn Ser Ala Met Asn Pro Ile Leu Tyr Asn Leu Met
580 585 590
Ser Ser Lys Phe Arg Ser Gly Phe Trp Arg Leu Leu Leu Thr Cys Leu

```
        595                 600                 605
Gly Gln Arg Pro His His His His Arg His His Tyr His Gln Arg Gln
    610                 615                 620
His Pro Thr Ala Gly Gly Ser Gly Arg Asn Ala Ser Thr Arg Gln Glu
625                 630                 635                 640
Gln Asp Ala Glu Glu Gly Ala Ala Leu Ala Gly Thr Thr Ser Ala Arg
                645                 650                 655
His Pro Arg Arg Thr Leu Arg Arg Glu Ala Thr Phe Leu Ile Asn Ser
            660                 665                 670
Ile Ser Thr Ser Ser Gly Thr Asp Arg Thr Thr Ser Ser Ser Ala Trp
        675                 680                 685
Arg Ser Asn Ser Leu Ser Ile Ser Gly Leu Ser Glu Arg Glu Arg Gly
    690                 695                 700
Ile Leu Gly Ala Ala Ile Ile Gly Thr Thr Ala Ala Thr Val Thr Thr
705                 710                 715                 720
Ala Cys Leu Gln Glu Arg Arg Ala Ser Lys Ile
                725                 730
```

<210> 40
<211> 3821
<212> DNA
<213> Drosophila

<400> 40

```
cagaaagtgt agggaatgta gggaaatgca aagcgagacc ctttggctgc tcactgactg 60
aaatatcacc agactcggct cggttctgcc gctcaagtgg cagtaaataa aagttgtgag 120
gagccttggc ggcatgagta ataatgaccg aaatcagagc accatcacca tcaccaacca 180
tcgctaccac cgcccaacgc gttgctcatt ttattaaaat aacaaaaata ttttgccgct 240
cgcagccttt ttacgatgcc gaaccgactg gcggcccatt tgttatttgg cggtcattgc 300
caccaaatgg tgtcctgtcg ctctttctcg tcctcgtcct cgtccttgtc cgttccactg 360
ggccaactcg tgtgtcgtgt gtgttttgac cagagttatg cagctgcgga tgggcaaaaa 420
tgttttattg ccagctgggg aatgccaaag tggcaagagg cagattgccc gatcttagct 480
gtttaaataa atttaagaga gagagagaga gagagagatc tggcacgaag cactccatca 540
ctggtccagc gatccgttaa gccatcgagg gtccttcgaa tgaaatgcca gcgctgacaa 600
ggcaaacact ggagatttgt ataataaggg gtgggtggtt gtgccgaaat ttggcactga 660
taggcaggac acttgtcagg cttggcgcaa acaaaactga catttatatt gatgtgtccc 720
tgtacccaat tagatgccca tcattatgag agcaggctat ctctgctctc gtgtccttcc 780
cccttctatt ttgcgcccgg gaagtcatat atcgaacatt aggctctgtt tgtgtgcaga 840
gcaatttggc gggaagtggc agcaggagga cgcacatgat attagtaata atgtaattat 900
atggtcataa acaaaataat aaatcaataa acaacacaca cgcacacgca ggcggcacac 960
gagtaaaagt tatcacgggg gggaaaaaaa tgaaggaaaa atgggaggca gagcagagga 1020
agtgaagaaa agggaaccgg aaaggaaagc cattcatttg gccccgcgac tcaagtgccg 1080
ttgtctacgt ttgcgaatgt gtgtggttgt gtgcgtggca ggacttcctg ttgccacatg 1140
gacggaaacg gcgccataat tatgccataa acccaaccat aaaggacatt tttatagatg 1200
tgcccgagtt aatccctttg ttgttccact cgctaccgca cccatctctc atgtcctttt 1260
tatccgcaca tcctttcagc ttactccacc gccgccctgc tggccatcgt gttcgccatt 1320
cccttcggta tccgcatgct ggtccacaag gatcgaggac agtgggagga gttcggccca 1380
gccttctaca ccgcccacct ggagctttat ctgggcaacg gctgtttggg tgagtctagg 1440
attttggatt tgcaacagga ggagctttca agggtcgaag ctcttctaaa aggttcaaaa 1500
gttggactca gctttaagtt cactcggata tactaaaaat atatataaaa caacacgtta 1560
tgaaactgtt tatgtaccat aaatggagcg taatgtgctt gatttaaagg aacagatgtc 1620
aacagtgctt aaattgtaat aaaaagctat ttgttttagg agctagccag acattattag 1680
tatggactga taagcatatg ttagaagata atatttcata catttagata cgcctataat 1740
ttcataataa taaatcataa aatttattcc gttgggatca tttaatcctt gaacagatat 1800
ctgttatttg ataatattta caaggaatat tatcccattt gcactcctag gcgttggcgt 1860
aatgatgtta ctagtgctga ccattgaacg ttacgtgtcc gtttgccatc caggatttgc 1920
tcgaccagtg atgggaccac ctgggtaagt cgcattgaaa gtctgtagat ggtctattta 1980
attgatccaa cctagcgttg tggtattcct cacctgcttg gccacggtaa tcgtctatct 2040
accgagcatc tttcgtggcg aactgatcaa atgcatcctt ggatcgagtg acgtatatgt 2100
```

atatttgcga cgcgacaaca ctatctacca gcagaccatc ttctatcgcg tctacaagat 2160
catgctggag gtgatcttta aactggtgcc cactttggtg atcggtggcc tcaacatgcg 2220
catcatgatg gtctacaggc ggacctgcga gcgccgccgc aagatggtgc tcagtcgtcc 2280
gcatgcccag ggacatggtc acggacatgg gcacggacat ggtcatggac atggacatgc 2340
acatggacac ggctatttga aggacgacga tccgcgaaag ttcgccgagg agcggcgctt 2400
gttcctgctg ctcggcagca catcgatttt gttcctggtc tgcgtctcgc cgatggccat 2460
tctccacatg acgatcgcct cggaggtcta tccgagcttc cccttccagg tgttccgtgc 2520
cagcgccaat ctgctggagt tgattaacta ctcgctgacc ttctacatct actgcctgtt 2580
cagcgaggat ttccgcaaca ccctggtgag gacgatcaag tggccctggt tgaagggcaa 2640
gttctgccac caggccgagc acgaggtgag tgccagtcca ccggccacag ccggaaccgt 2700
agcggtagcc ggaaccggaa atggccacgt ttctattttt catccggcaa taccggcgct 2760
gactctcaca cccgcagaac ccgacgaacg gccccggtgt gccaatggcg tgcttcacta 2820
aggtggacag gggccatcaa aagcaccaca taactaccac ctcgggcctg ggacgatcga 2880
gtagcattta gacaatgcct accacccaac aacaacacca ccacaaccca tccgctactc 2940
gagatccgaa ctccgaacac cgagaaccct cccagcccag agtaacccca ccaccaccac 3000
ccaactccaa aaccgtcatc accacttcca cttacccccc aggtcgtcga ggttaaaggg 3060
tgccctttct ctcgcttgag atcctgcgac atttgaacgt gttcaagtat cgatgaacga 3120
aaaatgatat acgataaaca actaacgatt aacgataatt gatattcagc gccacgcatt 3180
gcgtatataa gtgtctgtgt gtgttcctta aacatatata tttaaatata tggataaaca 3240
ttttttttt tttagcacta gaattaggta cttaaaaacc aaacttaatc gaaaacacat 3300
ccgtaattgg tgaaataatt gatatcgata ggtaaaacca gttcgacttt cgatcagcat 3360
ttgcgccgaa tgatttcacc acaatgtagc ttgcggaatg tgtatgtcgc aggatcaaga 3420
tcctcaagcc acacgtatag atgcagatat gatagtttta tcaatggata gattgcttgc 3480
cgtattgata accaaagagc tcgtttactc tagcattaga tgcagacaga aacaggaaca 3540
gcgacaggac aacgacagac agacatattg acatatagac aggtatagac aggtgagatg 3600
gaacaatagc gagagagaga gagaaagagt tttacgatga taatatgcag acatacaata 3660
cttatatata cacttgcata tgcctatatg tgaatgatat ttatgtaaac gacgtattta 3720
tgtatacacg cgtaacttta agcgaatata tttatatatg catatatgta tctattgatt 3780
atagcattta agtgcgcgca agaggaatcc ggattccaac c 3821

<210> 41
<211> 1158
<212> DNA
<213> Drosophila

<400> 41
atgggaggca gagcagagga agtgaagaaa agggaaccgg aaaggaaagc cattcatttg 60
gccccgcgac tcaagtgccg ttgtctacct tactccaccg ccgccctgct ggccatcgtg 120
ttcgccattc ccttcggtat ccgcatgctg gtccacaagg atcgaggaca gtgggaggag 180
ttcggcccag ccttctacac cgcccacctg gagctttatc tgggcaacgg ctgtttgggc 240
gttggcgtaa tgatgttact agtgctgacc attgaacgtt acgtgtccgt ttgccatcca 300
ggatttgctc gaccagtgat gggaccacct ggcgttgtgg tattcctcac ctgcttggcc 360
acggtaatcg tctatctacc gagcatcttt cgtggcgaac tgatcaaatg catccttgga 420
tcgagtgacg tatatgtata tttgcgacgc gacaacacta tctaccagca gaccatcttc 480
tatcgcgtct acaagatcat gctggaggtg atctttaaac tggtgcccac tttggtgatc 540
ggtggcctca acatgcgcat catgatggtc tacaggcgga cctgcgagcg ccgccgcaag 600
atggtgctca gtcgtccgca tgcccaggga catggtcacg gacatgggca cggacatggt 660
catggacatg gacatgcaca tggacacggc tatttgaagg acgacgatcc gcgaaagttc 720
gccgaggagc ggcgcttgtt cctgctgctc ggcagcacat cgattttgtt cctggtctgc 780
gtctcgccga tggccattct ccacatgacg atcgcctcgg aggtctatcc gagcttcccc 840
ttccaggtgt tccgtgccag cgccaatctg ctggagttga ttaactactc gctgaccttc 900
tacatctact gcctgttcag cgaggatttc cgcaacaccc tggtgaggac gatcaagtgg 960
ccctggttga agggcaagtt ctgccaccag gccgagcacg aggtgagtgc cagtccaccg 1020
gccacagccg gaaccgtagc ggtagccgga accggaaatg gccacgtttc tatttttcat 1080
ccggcaatac cggcgctgac tctcacaccc gcagaacccg acgaacggcc ccggtgtgcc 1140
aatggcgtgc ttcactaa 1158

<210> 42
<211> 385
<212> PRT

<213> Drosophila

<400> 42
```
Met Gly Gly Arg Ala Glu Glu Val Lys Lys Arg Glu Pro Glu Arg Lys
 1               5                  10                  15
Ala Ile His Leu Ala Pro Arg Leu Lys Cys Arg Cys Leu Pro Tyr Ser
            20                  25                  30
Thr Ala Ala Leu Leu Ala Ile Val Phe Ala Ile Pro Phe Gly Ile Arg
        35                  40                  45
Met Leu Val His Lys Asp Arg Gly Gln Trp Glu Glu Phe Gly Pro Ala
    50                  55                  60
Phe Tyr Thr Ala His Leu Glu Leu Tyr Leu Gly Asn Gly Cys Leu Gly
65                  70                  75                  80
Val Gly Val Met Met Leu Leu Val Leu Thr Ile Glu Arg Tyr Val Ser
                85                  90                  95
Val Cys His Pro Gly Phe Ala Arg Pro Val Met Gly Pro Pro Gly Val
            100                 105                 110
Val Val Phe Leu Thr Cys Leu Ala Thr Val Ile Val Tyr Leu Pro Ser
        115                 120                 125
Ile Phe Arg Gly Glu Leu Ile Lys Cys Ile Leu Gly Ser Ser Asp Val
    130                 135                 140
Tyr Val Tyr Leu Arg Arg Asp Asn Thr Ile Tyr Gln Gln Thr Ile Phe
145                 150                 155                 160
Tyr Arg Val Tyr Lys Ile Met Leu Glu Val Ile Phe Lys Leu Val Pro
                165                 170                 175
Thr Leu Val Ile Gly Gly Leu Asn Met Arg Ile Met Met Val Tyr Arg
            180                 185                 190
Arg Thr Cys Glu Arg Arg Arg Lys Met Val Leu Ser Arg Pro His Ala
        195                 200                 205
Gln Gly His Gly His Gly His Gly His Gly His Gly His Gly His Gly
    210                 215                 220
His Ala His Gly His Gly Tyr Leu Lys Asp Asp Asp Pro Arg Lys Phe
225                 230                 235                 240
Ala Glu Glu Arg Arg Leu Phe Leu Leu Leu Gly Ser Thr Ser Ile Leu
                245                 250                 255
Phe Leu Val Cys Val Ser Pro Met Ala Ile Leu His Met Thr Ile Ala
            260                 265                 270
Ser Glu Val Tyr Pro Ser Phe Pro Phe Gln Val Phe Arg Ala Ser Ala
        275                 280                 285
Asn Leu Leu Glu Leu Ile Asn Tyr Ser Leu Thr Phe Tyr Ile Tyr Cys
    290                 295                 300
Leu Phe Ser Glu Asp Phe Arg Asn Thr Leu Val Arg Thr Ile Lys Trp
305                 310                 315                 320
Pro Trp Leu Lys Gly Lys Phe Cys His Gln Ala Glu His Glu Val Ser
                325                 330                 335
Ala Ser Pro Pro Ala Thr Ala Gly Thr Val Ala Val Ala Gly Thr Gly
            340                 345                 350
Asn Gly His Val Ser Ile Phe His Pro Ala Ile Pro Ala Leu Thr Leu
        355                 360                 365
Thr Pro Ala Glu Pro Asp Glu Arg Pro Arg Cys Ala Asn Gly Val Leu
    370                 375                 380
His
385
```

<210> 43
<211> 14111
<212> DNA
<213> Drosophila

<400> 43
aataatccac cttcgaaggt aaaccatctt aattatgctt taaaattgct ggattatata 60
aacccctaag gtagttcatc ccacttgttc aactgaaatc gacccacaca aatggtggtt 120
ggtatctatc taccagagtt aggagctcgt ttaacccaat attgcctatt cgtagagatc 180
cgcattcctg cgaggagaag tcgaactgcc gtagctgagc tgcgagtcgg tcttcttcag 240
aatggaaacg tggtgggtgg gtgcatcacc cgtccggttg gccatcagcg agtcctgttt 300
gtggagcagg ccgcgtcgct ggtcggcgaa ggagatgacc gagtcctgct tcaccagttg 360
atggcgcctt tgctctggtc gctggtccac aaacgagatc accgagtcct gctttacgaa 420
cttctgccgc ttatcgagca gagagctcgg tgtggcgccc tgtcgccggc ccacatccat 480
ggaaacgcta tccctgtcca ccagtgtgtc gttgagaccg tgtcctgccg acctgcgtgc 540
ctccatgtag gacacaattg agtcctgctt ggacagggag accttggagc cattgctgtc 600
ctgctggcag aggggtcgcc ggatcccaga agcatcactg caggagctct cctgcttgat 660
gaggaactga cgcgatccgt gagagtcctg cttaagcagc atctgctgcg acaggttgct 720
gtcctgcttg agcatcatgt tgtggcgccg cagcgagttg tgctcgtaga tgtaggacac 780
cgcgttgccc ttgctgctcg gactcttcct cgagctgttc tggtggtagt gcgatgactt 840
ccgggtgctg gtggtgctgg aggcgatggc cacgttgtcg tagtacctag gtggttaagg 900
tggaaaaagg gtttttatcg aagggaacac gatttcttac aaatcttcat atcaattttg 960
aaggccttca catattttat caatatattt gaggttagaa attaactaac ctgaggcgtc 1020
cgcagcagct gcgcgacttg atgatggccg cgaatccgcg tcggtacttc ttgttcactg 1080
agtagagtat ggggttgatg cagctgttcg aggagcccag ccactgggcc accggcatca 1140
ccttcttgag gatctcgaac tcctcctgcg agatgtccga gccgaatttg atccgcgcaa 1200
agatgacata gagcggcagc caggacaaaa caaacaggat gaccacggcc accaacatct 1260
tgatgacctt caccttgctc ttctgctgca tgcggtccat ttgcgcgtcc ttggactcgc 1320
cggggatgga cctcgtggag accttgatcc agatgagcac gtagcagagc gtgatcaggg 1380
acatgggcag caggtagcag gccaccaggt tggctagcag gaagtagagg ttcccatccg 1440
tgcccggggg ccacacctcc tggcagagga actgcggctg cgagtaggcc gagaccaggg 1500
cgtcggagaa gacctcctcg gcgggcacca ggtcgaagaa gagcagccaa ggaatggtgg 1560
tcaccagcgc gattacccat atgccgatta tcatgatgcg ggcacgtcgc tttgtcatct 1620
gcttcagggg ccaccagatg gcaatgaacc tgcaagaggg gagtcccaac aaaggggatt 1680
tttaccaagt agtcaaagaa ggagtcttcg taggagcagc agcagcagca ggagctccgg 1740
cttcagatgc cactgacggc tttgagccgt caaggataat ggcaatcatc aagctgtcga 1800
gtcgagtcga ggcgacgatg acgatgacga caccgtttca actaggagga gcagctcggc 1860
atgcaaatat ctccgctgat taattcgcac acgcaggatg ccgatgccag cagccatcct 1920
taagcctttt tctgctgagt gttcgctttc caatctcgac tattgaaccc gcctgctact 1980
cctgctcctg ctcctccgcc ccaaaaaagc ggcgctgcgg atcctgtcat tttttattct 2040
gcttgccata atgaattgta atttgaataa gagctgcgag gtagtgagac gagtcggtat 2100
ggggccaaca acgagttcag catttttcat ttagcttagt gcagccggct cattgtggcg 2160
cccgctcagt ggtcatagac aagacttgtc ccggacgaac gaaactggtt gtctgctcct 2220
ggcccgaaac catgtaatta cgtgcattgt gttcgtattt tcgacttgga ccttgggttg 2280
tcgcctgcgt ctgtccgcca aaacatttgg cagataacac agcggtaatt ccattagagg 2340
gcggccatgc caggccacta ccccagaatc ctattcctca ccgactaaaa atagccagct 2400
taacaatgtg gccgagaccc tcatctctgc gaaaaaaaaa aggtgggccg cctgggcgac 2460
caaacaaagc tgtccgcaaa gccaataatc aaattgtcat aacgtttatt tataggagca 2520
aaccaattgc atttgcataa tttgcccatg cgacttctat ggtcataaat caggcgagag 2580
cttaattgag ttctttttgg gtaaaaacag cggagtcgga aaaagtaatc agcagttttc 2640
gctggtgcag actaattcaa actaaacctc taatttgatg ttgaaatttt gttaaatcgt 2700
attaactaca cttttgatga tgtaaaccat tttatgagct ttaaaaacta aaatttctct 2760
tatgagttct cttatacgat ttatgtatga tttgtttcac aaataaactc tccttataga 2820
ttcaattgct gtcggaaaat taaattatta atattatttt ggtaatcacg ggtaattgtt 2880
gcccatcaat tccgttgagt accatatgcg taaatctcta aggacgtcag ttcagtattg 2940
ttagcatatt tattaagaaa aatcttgcca caatcaactt aaacaaagta gttaatactc 3000
gcgcttttaa agcatttaat aagccgcaag tgtattaatt taatatttgc cacacaatcc 3060
tgtgtttctg tctagatcca cgtggtttta tggctcctat ttcgcttcct actgcggtta 3120
cccatacgca ctgttgacat ttgcgacgca aagaagacac aaaagccgca atgccagcac 3180
cgtaaagaca caactatgtc tataagcgag tgtgctttct tatagacttt atgattataa 3240
gcaatctagt tatatacata tacgtattca gacaagaaaa gaccctaatt aaatgacagt 3300
ggcataagta taaatataag tagaagtaga agtataagca agtgataagc taagtcgaag 3360
gggcattacg ggctctaaga ccgtggagag atttaagttt catgttacag ataataaatc 3420
ttaaatttcg attcagttta cgatggctcg taaaaggaga gaataacagc aaaacaactg 3480
gataaacgga catcaatgac gctggactcg gacttcgact gggacttgga agtggaagtg 3540

```
gaattggaag tggatgagga tgtggatgtg gatttggact cggttgggtt ggcgttaagc 3600
gtgggcggag gctgacatgc gggcctgtgg gtgaccaaac gatcaagtga ggcgctatgc 3660
ctttgtcccg ctggctgaaa aatgacagtg gcaaaaacag aagcttctgg ccaaaactcc 3720
tcctctgatt gtggccagaa attgatataa atgaaagtgc acataaagcg acgggcgtct 3780
tattgtcctt gcggctatgg ctagcacaat ttgagggact tcagcatcac ttgataaatc 3840
tgttgaccgc gacatttgcc gaatttattg ctttttttata ctttatgggc caggaaacgc 3900
agggctgaca catcgattcc tggcggccaa tcatcaatgg tatcaatatt cgcggtatcg 3960
cggccaacaa cattctaagg caattgaaat gcacataaaa tgagaaacat gaaattgatt 4020
gtaagcagcg cctaatacgg acaaggtgag ccatagaagc atatgatgat cttgatcttg 4080
agacgactgc tccactgact ctcctacata actgcctaac tggctaactg ccttactgac 4140
cagttttccc cgcagctgat tgcattgagt ggcgctcact ttcggttcgg tcgcctggtg 4200
caaatccact ataaaagccg cactcggagc tcaagttgtc accagttgac aatttgaaat 4260
cattagccta gctacctttc ctacttcagt tatcccgcca aaatccaacg cctcaaaatg 4320
ttcaaagtag tgagtatcat tgtgcagaag aagttgtaac cagtttgaat ccaagcatcg 4380
tatcattatg atattatatc agtttgctct atagcattta aaaataaatg caataaatta 4440
agtttatatt acaactcata tgagtcacct agtttgttgt taggattggt cttatcttac 4500
tcagaatgag taatatatgc tttaaagtct gcgacaacta accaaattca ctgtggagtg 4560
gcagccacct gctgctccgt ctaatgccac taatcaggct aattcgtgaa ctctgatccg 4620
ccgcagctgt tcgtgctcgc cgccttcgtc gcctctcagg ctatcgccca tcccggcgtg 4680
gtggccgtgg cacccgtggt ggcgcacccg gcggtggtcc acacgcccat catccatcat 4740
ggagcccact cggtgcactc gtaaggccat tcatgtccaa tctgcttatc accaccgaac 4800
aactcacccc tttcattttc agccatgttg ttcatcatcc ggcagccgtc aaggtcatca 4860
cacccgtcgt ccacaagccc gtggtggcgg tgcatgctgt caggccggtg gttccattgg 4920
ttccagtgca tcatgccgcc cccgccgtcg tcgtgcatca ttaagtgcag cgaatacggt 4980
tcactttcaa tacctttccc gctcccgatc ccgttcccga tccccaccag acccagacca 5040
atcccttcct gagcaccata accacgatcc cacaataaaa gttcatagcc acgtaattgg 5100
agcgagtgtt ttcaacttac ctgtccaggg aaacggcaat caggctgtag accgaggctg 5160
ccacggagac accttgtatg tagggcacaa acttgcacat cagccatccg agcatccagg 5220
ctgcgagttc cagcagttcg cgaaacatga aaggggggaaa agtgcgggag gtgagtaaat 5280
attggtcata aacaacgcac aaattatgct gagtaacttt cggtcacatt gtccttgctg 5340
gtgtcctagc catacactga aggaaatgtg agggcctatg gcaacaaaaa aagaagaaga 5400
tgcagatttt tctaaaatac cccaaggaca ctttgctcga aaactttata aatatcctaa 5460
ccacgaaaat caaagcaaat taatgtggtt aaccgaccga aggatgctta ttgttctcat 5520
gggaaatatt cgaaatagta gatggtgtat cactttaaat acaaaccgta agtcatcttg 5580
actgcttagc taagaaaaga gtgaaaatat tataaaacta ataattctcc gctaagtatt 5640
ataatatatt tcatatctta cttttattct tgaataagta cagttttaca gctggggatc 5700
attgagttct ttatagcttt actgcttttc cccgagtgca ctggtagtgg ttatgctcat 5760
ggttgtggtg atggtggtgg tggtgatagt ggtagtcatg gacagttggg ttcgagatct 5820
ggttagcgtt caatgcttct tgtcattgcc gccgcccacg ttcccgtcgt tgtttgctga 5880
ctgagattta tgtgctgcat atcattcata atttggtgac aagtctcaga ttgcgcccag 5940
ccatcgtagt gggtggggca gaatgtggag tcctgggggg cagagcattg tttgcccaga 6000
cccagactca aagccgcggt ggaccgcagg actgcctaac ttaatttgta ttccttttac 6060
tccattttgt ggctggatga tgatggtaaa tggaatacac gggccctact tgcgcatagt 6120
tgagcgttat gaattttatg ctaagaggca taactcaccc agacacaaaa ggttataaag 6180
catacataca tgggcgtgca atgggagggg caacttccca gggttgatga cagtcgggcg 6240
gagaaaggcg cccggttcat gacatatcga atggcaatta aacaagtgca agtgcggccc 6300
acacatcatt agcagttagc tgtttggcag tttggcaatt tgcagaagga gtttaccgag 6360
ttcaatgctc gaccaggaga cgcctcctgc gtcatggggc tcctaatttc ctgtctcgtt 6420
cgcctgtctt atctctattt gttgtcgctt tttattatgc agacgcaccg taagccgcca 6480
aggcagccca ataaaagcta actttaatat gatttggcac ataaaaatat aaatattatt 6540
ggcgctgtgg cgcttcgtgc aagtgaactc tgtggctgtg gacaggtttt ggttctggac 6600
aggatcatgt gtgtctgctg ggaggacata cacatcctcc cgcctggcaa acaatcaaca 6660
gtgtgtacac aggacagcag accgaattgg tttttcggtt ttattcctcc gatttgaaat 6720
ggctttcggg cagacctgca agtgcaacag actggatccg aggctattaa tgcgagtgta 6780
tttggccact cacttgatat tatgctgata agggtgaact ttgatatata gctgtctgcg 6840
ggagctcttg actcggcaaa ctttatgctt tcaagttgtt tgaaaacaat ttaaattcaa 6900
attcggttcc gaaatgcttc atcacttgcg ggacaccgga ttagcagtgc cactgggtga 6960
actttttttt ttttttaaat acatgttctt ccaacttccc ttcggcaaat aaactatcta 7020
attgtgtatg caattctatt aggacaggca aacgtgactc ggttggtttg tctgggatta 7080
aagtttgcat cagtgcactt tggcagtcaa gtgtcttaac caagtttaca ttgtggttag 7140
```

cttagcgcag tcaagtgacc acttgataat cccgaacaaa caatggtttt cagttcatat 7200
ggaaagtatt ttgacaggct gaggcattca gcacggttca acttattaaa ttgtccaact 7260
cgctgatttg gatatttaat ataaaaatgg aaaaaccaca ttttacagaa ccaacgagaa 7320
aagctgacat ttaaaattgg tttttgttaa cataaaacga ttgagaagga ataatacata 7380
attaatgaac cgcaaaaacc gcattatata taaaaaaaat caaataggtc aagggttatg 7440
ccagtagtta tttatagata ccacaagctt aggtaagaca tgattatttc aaggaacacc 7500
gagtcattcg aaaccaatct ttaattaacc ccaaccttaa ggccacatta atgtgttgta 7560
caaacataat tgtagctacc gaaataacct aattgccttg cgtgacgggg ctaacccgaa 7620
cctccttgtg gcattaccca tttgaggact caagtgaaac caaacactcc acagtttgct 7680
cggccaacat ccagcctccc cgtcaaggtg agaatgggat gggtgtttga tccgacttgg 7740
agtgcgtttg acatggcaac atatggcaaa cagacgcctc catttatacg cgaaagctgg 7800
tgctcaatat tgatcaacta gaggagataa taccaatcga tactcaggct tatcgccagc 7860
cttcgattgt gggtggtggg tggtgggctt gtgcgcccca tcagccctta agaaaatgtc 7920
aagtgcgccc taagcctttt attttttagc ctcttaccct ccttacaacg agagcctttg 7980
tttttcggtt tcggtatata tgacttttag aaatgtgatt caactgaaaa tatttacctt 8040
atttgataat ttgcctgcag aattaaattt aacgcggtta tttatgaata atgtcccctt 8100
atttggaaac aaatttttct caaggacaac tttaaaacgc gttatctttt acagcaacaa 8160
taaacatttt ctctactgtg ctttatgagg ctttaatagc gccaacaaag ttttgttaca 8220
tttttatccc caataccgtg ggtgattgtt cctatgtagg caaagtggat tggattattc 8280
cgggcattgc ggattcgttt cggtggcggc attgtcatgc acggagtttg atcgtaaagt 8340
taatgaaata atcaggcgtg acaaatccgc aggtttggac ttgggagtgg gaacgggcgt 8400
ggattgcagt ggagtggagt ggagtggagt ggccatggaa aaggcataac ataacacttc 8460
tggcaaatgc gttattgtta tgattatcgc cgcggcgtca cgtcctgcgg cgacttctgt 8520
ggcagctctt cgggaaaatt gaataccaca gctagttcgc agagaaatgc gcctgggacg 8580
cggagtggga ataggaattg gaatctggga atcggtggtc ctcgatgtga gatacgggat 8640
tgcgggagta aagaagctgc agatcttagg tggagggctc agaggttgga tgccatttgc 8700
cgacggcacg tagagactaa attaaaatac aattggcgag ctgaaaaatg acttcatcag 8760
ggatagggaa acagaaatgc gggagagtct gcctccaagt tacgcatacg ccgcgttcgc 8820
ccggcacaaa tcacgtggcc ccctaatgaa caaactggca acgtgactga aagatttagc 8880
caacgggaga tgctgatcta aaaagggccg taatgataac gattccccat tgataaaata 8940
gcacaacaat ggccataaag cgccatctga gacaattgcc gttgctgttg caaggatatc 9000
cttcgtctca aggcggcatc gttgctggga aatcaattca tttcgcattt tattggattt 9060
caaagcacac acataaataa gggccagcgc cacattccga ccaacaacca ccacccacca 9120
cccaccaccc acaggccacc agccaggaac cacccacaaa gagtacacgg cccaacccaa 9180
tacaggtgct cccacacccc acgcatacac acacacacac accctcaacc acacacactg 9240
gcacactcgc aagtgtgtgt gtgcataacc ctgactatta ggtttgccaa gcgaaatatt 9300
taggtcaatt catccagccg caaagagcct tgaatgcaaa caaccttcca ttgcaattct 9360
gaggcttgaa ggtgtgctta gataatggcg ctgccagcgg cgatggccgt gggcgtcgct 9420
gccagctgcg actgcggcgg cgacgtcgtc tttgacggct ggaatttgat ttcatataat 9480
tcattttggt ggcaaaacaa cccgcccact ttgccactcg ctgggggcga gtagtgggaa 9540
ctatcagaca gcgatgtgaa atgctcgatg tacaaatatg ctaaacttcc gaactcccac 9600
gcactgtgct acacagggag aaaaggtttt gttagaggag agcaaaggat taagtggtat 9660
gactttacac tgctaatttg attatgaaga atattcgtaa gagttatagc atttcatcac 9720
ctgacgaagt tcaaaaacat ttaagtttaa agccaaaaag tgtatatttc tgaaaacgtg 9780
tggattgttt aatgaaaatg acgtgttaaa cttttgtttg taaataatac gctacacatg 9840
tgtttagttg gcaacacatg ttttttcccc tgcattagtg ctatttcacc gttgccgctg 9900
tgaaaccgtt gccttcgttt taacttcagc aggttgtcga aagtgaatag gaatgtaagt 9960
ggggttgaga gttgtgggga tgaagaggtg aagctgaaat gagccacccc aaaagcccca 10020
atgccagcgg gagggctgca atttgttgtg actggcgttt ttgctaggca tcggcggctc 10080
agctgaagct aagacatgcc aagcgaaaat tgcggagtgg gcggccaaaa gggtgcccca 10140
tccaaacaac catacaacca tctagcccac ccaatctacc atccaccgcc agccacccac 10200
caccccagt gacccagtaa gagctgtcgg ccctgcgccc agcttgaatt tgtgaaatta 10260
ggtcagtttt gtgtcgatga tcctttgggc caccccgaaa agcagccacc aatgcaccta 10320
cacatgcaca gtagcgagca aacacatggg cgtgagagcg ggcgagagag agagagcgga 10380
gagagtgcgc agtgacacgg acacaaaaac ggagacacgg acaccgatac agatacacaa 10440
tgcaggcgga ggtaaactca ttcacttttc gccactcgct tcattcgttt ccagcgattc 10500
gcactcggca gccttcgttt actacatttc ccccacagtg ggcttctctg ggggagtgta 10560
ttggtatggg gtgggatcga gtgcagtgcc gagtctaaga gtagagctgc attggaaaat 10620
aaatctgata agtcaactgc atctgcatat ttgttaccag tttgtaatat taatattcag 10680
gttaataaat aatctgaggg gaagaggaac taaaactatt tttattatta ttattattaa 10740

```
tttaatttaa tttaatatta ttattattaa ttattttaaa taattttaat tattttcata 10800
tacttgtggc aaagaaatgt agtaacacaa tttgttaaac tatttcctat taatcgaaca 10860
tattgtttta cgagtatgtt gacaactctt attttgccag cgtactcgct gtatagcaac 10920
aagaaaagga gagttctcac agggctgagc cacttaattg gcaacacttg gcacaacctg 10980
cgagcctacg acctggcatt gggatttggg actagggacc tggcactggg attggggtag 11040
ccatcgaagg caaattgcag tgaaggatgg agcttgcatg gcgtgctcca caatcggaat 11100
ttgatttacc aaggatttcg atcatttttc catgtcagcg gcgatttgca tagacatttg 11160
ccaacctgct caccttgtcg tccaggtctg cgactatcaa tttggatgtc agtccacaaa 11220
gttcgtttcg aatgcaagcg gggaagttgc ggagggggaaa ccaaccaatt gggggtgtaa 11280
gacggcagac agggacagat agatagatag atagatggac ggatggatag gcatgtcagg 11340
gaaatttatt taaattaaag catacaatta tgcgtagcct cgtttatcgg gtcgtttgtg 11400
ggtcgtctct gggatttggt ggaattggcc tctaattaaa tggacacttg ttctttggtc 11460
catgcggatt gtcgtttcca agaaggtaag ggctcccatg gagagagcac aaattaattt 11520
aattgagtgc gaacatccgt tgctacggaa atctgacaca tttctggtaa agtacgctga 11580
atgaattctt cttttttttct ttgctgcgta ctgtgtgaga tgtagttcct tgatttaatg 11640
aataactggc ttttgtcttg gcagtctatt ttaaaagtta tcttgaccat tgcagaaagc 11700
tattaaattt tgatactaaa gcttggcatt ttccagaaca actttaatga tgctgcatcg 11760
ttatgttctt cttgtgtatg tagtaaaaat tcggaagaag ccagggagtc tggtatataa 11820
aatgcagctt acaaataaac ggaaacacag tatgtctgtc aaaagtaaat ataaccgagt 11880
atgaataaaa tcccataatt tattattttc caccagaaat atgagcactt tcgtttggac 11940
atttatgccg acattatgtc taatgagggt caattacact tgcccaaaca tatcgggcac 12000
attacgcata cgcagcgtaa gcacgcagga gcaacaactt tcgtcagtgc catttggcag 12060
ggctttatct aaagcggcga ttaattgcag tacataaatc gggctaatat gcagaacaaa 12120
caaaaaaaaa cttcggcttc cccctagcag ctaacgtggc ccccagtggc gaataaatct 12180
cgtcggcatc ttgcggaatt aaagagatac ttacgcacaa agatgttgcc aatcagtgtg 12240
gctggcaggc agaagacaat gaccaggatg tccgcaatgg ccttttcgtt ggagaggggg 12300
gaatgtaaat tttatttcaa attaatgggg ttcatccacg agatttacgg ggtccaccca 12360
ccagattcac aatgaagtag ttcgtcaccg tccgcatgcg aggagcccgc aggacgacgg 12420
cgatcacgaa gctattcccg atgagtccca ccaggaagac gacgatgtag gccacgcagt 12480
agaccatgga catggccagc gagtgacggt agagtgggtc gaatccgaag tccacgctcc 12540
cgtccagctg ctcccaccgg agggcactgt cgttgtacga cttggtggcc gcagccgccg 12600
agatcaagct ggtggtcgcc ggcgataggt tcagcaagtt ggaatagtag gtggatgaag 12660
tggcgtctgc gatggccagc gacttgcccg ccagggatcc tctcgcagat gctgttgtgg 12720
tggccgccaa ggtggctagc agcgaactgg ccgccggagc caccgtggcc gtggccgtag 12780
tgagaaccgt gcccagcaga gtgtccgtaa tggcgctgtt attcagttta atggcactcg 12840
atgagcccaa ggagctgaca gcagtcccac cgctaccact agtgccattc aattcgttga 12900
aactcaagag gttctgcaaa tcgacgaaat tatgcgtgaa attattgtaa tccccggcga 12960
gggcggcggc ttcctccatt gtctccgcga ccattttccc cggcgccatg gtcggaatgg 13020
gcatgggtac actagtcgtc gtactgggaa cgtcgccact tgtatggctt ttatgctttc 13080
gcttgcgaat ccgcccgctg gctgccatca tcaaatcatc cggcgtcttg attgggcgga 13140
tggcactttc actggcgcct cgctttcgct atcactatct gtagacgtat ctcgaatctt 13200
gaatctggta tctcgcttgt tttagaaaag gtctataaaa atggcatcgt cacttggcat 13260
ttcgttggct tgtgtaggat atgcaggaag tttcgtggct ttgggttaga gaacagtttt 13320
taattaacac ccatgtccat acaatatctc gtttatccgg ggattgtttt gcgtttgccc 13380
cgaagttgca caataaatgc agttttctag aaactcgctt ctagcgccca aggacccttc 13440
cgttttcgtt tttttttttt ttgattgaaa tcacctacaa aagtttgttg ccgtttcggc 13500
tgcgaatccg agtgcggaat gcaatgcatt gcctgggtcg ttcgtcagct ggttgacgtt 13560
gccttggctg gcattacgaa aaagtcggcg gcccaaggcc tcgtgtgcat atttcagata 13620
aggccagatg ctgggtgcct ggataaatgc agctagaggg gattggggat cgggcattgg 13680
gcattgggca ttgggcattg ggcatcggac atcgggcaaa ggagccgaga gcatcggctc 13740
agggacatgg tcaccgccac agttgcaaca catgcgcgtc gctaattaat tttaatgccc 13800
gttctgccat tgtttggctg ttgtttttgg cacaattgaa aagccagcga gcgtgcactt 13860
tgacctcaaa aaatcgacac caaacatcga acaaacgagc aattgagcaa ctcaatagca 13920
gcgcataaaa attccattaa gcacaactgc caagcgcacg cagcactctt cgattgagtg 13980
ccaaaaaatt gccagtaagc attgaaacaa ttggtaacaa acgaagccag tgcgacattc 14040
ttttgtgcgg ttccatggga gtgtccacat cagacaaact ggcgactaag ggcccctcgg 14100
cacacaattc a                                                       14111
```

<210> 44
<211> 1525
<212> DNA
<213> Drosophila

<400> 44
atgatggcag ccagcgggcg gattcgcaag cgaaagcata aaagccatac aagtggcgac 60
gttcccagta cgacgactag tgtacccatg cccattccga ccatggcgcc ggggaaaatg 120
gtcgcggaga caatggagga agccgccgcc ctcgccgggg attacaataa tttcacgcat 180
aatttcgtcg atttgcagaa cctcttgagt ttcaacgaat tgaatggcac tagtggtagc 240
ggtgggactg ctgtcagctc cttgggctca tcgagtgcca ttaaactgaa taacagcgcc 300
attacggaca ctctgctggg cacggttctc actacggcca cggccacggt ggctccggcg 360
gccagttcgc tgctagccac cttggcggcc accacaacag catctgcgag aggatccctg 420
gcgggcaagt cgctggccat cgcagacgcc acttcatcca cctactattc caacttgctg 480
aacctatcgc cggcgaccac cagcttgatc tcggcggctg cggccaccaa gtcgtacaac 540
gacagtgccc tccggtggga gcagctggac gggagcgtgg acttcggatt cgacccactc 600
taccgtcact cgctggccat gtccatggtc tactgcgtgg cctacatcgt cgtcttcctg 660
gtgggactca tcgggaatag cttcgtgatc gccgtcgtcc tgcgggctcc tcgcatgcgg 720
acggtgacga actacttcat tgtgaatctg gccattgcgg acatcctggt cattgtcttc 780
tgcctgccag ccacactgat tggcaacatc tttgtgcgta agtatctctt taattccgca 840
agatgccgac gagatttatt cgccactggg ggccacgttc attgccatct ggtggcccct 900
gaagcagatg acaaagcgac gtgcccgcat catgataatc ggcatatggg taatcgcgct 960
ggtgaccacc attccttggc tgctcttctt cgacctggtg cccgccgagg aggtcttctc 1020
cgacgccctg gtctcggcct actcgcagcc gcagttcctc tgccaggagg tgtggccccc 1080
gggcacggat gggaacctct acttcctgct agccaacctg gtggcctgct acctgctgcc 1140
catgtccctg atcacgctct gctacgtgct catctggatc aaggtctcca cgaggtccat 1200
ccccggcgag tccaaggacg cgcaaatgga ccgcatgcag cagaagagca aggtgaaggt 1260
catcaagatg ttggtggccg tggtcatcct gtttgttttg tcctggctgc cgctctatgt 1320
catctttgcg cggatcaaat tcggctcgga catctcgcag gaggagttcg agatcctcaa 1380
gaaggtgatg ccggtggccc agtggctggg ctcctcgaac agctgcatca accccatact 1440
ctactcagtg aacaagaagt accgacgcgg attcgcggcc atcatcaagt cgcgcagctg 1500
ctgcggacgc ctcaggttag ttaat 1525

<210> 45
<211> 507
<212> PRT
<213> Drosophila

<400> 45
Met Met Ala Ala Ser Gly Arg Ile Arg Lys Arg Lys His Lys Ser His
1 5 10 15
Thr Ser Gly Asp Val Pro Ser Thr Thr Thr Ser Val Pro Met Pro Ile
20 25 30
Pro Thr Met Ala Pro Gly Lys Met Val Ala Glu Thr Met Glu Glu Ala
35 40 45
Ala Ala Leu Ala Gly Asp Tyr Asn Asn Phe Thr His Asn Phe Val Asp
50 55 60
Leu Gln Asn Leu Leu Ser Phe Asn Glu Leu Asn Gly Thr Ser Gly Ser
65 70 75 80
Gly Gly Thr Ala Val Ser Ser Leu Gly Ser Ser Ser Ala Ile Lys Leu
85 90 95
Asn Asn Ser Ala Ile Thr Asp Thr Leu Leu Gly Thr Val Leu Thr Thr
100 105 110
Ala Thr Ala Thr Val Ala Pro Ala Ala Ser Ser Leu Leu Ala Thr Leu
115 120 125
Ala Ala Thr Thr Thr Ala Ser Ala Arg Gly Ser Leu Ala Gly Lys Ser
130 135 140
Leu Ala Ile Ala Asp Ala Thr Ser Ser Thr Tyr Tyr Ser Asn Leu Leu
145 150 155 160
Asn Leu Ser Pro Ala Thr Thr Ser Leu Ile Ser Ala Ala Ala Ala Thr

```
                165                 170                 175
Lys Ser Tyr Asn Asp Ser Ala Leu Arg Trp Glu Gln Leu Asp Gly Ser
            180                 185                 190
Val Asp Phe Gly Phe Asp Pro Leu Tyr Arg His Ser Leu Ala Met Ser
        195                 200                 205
Met Val Tyr Cys Val Ala Tyr Ile Val Val Phe Leu Val Gly Leu Ile
    210                 215                 220
Gly Asn Ser Phe Val Ile Ala Val Val Leu Arg Ala Pro Arg Met Arg
225                 230                 235                 240
Thr Val Thr Asn Tyr Phe Ile Val Asn Leu Ala Ile Ala Asp Ile Leu
                245                 250                 255
Val Ile Val Phe Cys Leu Pro Ala Thr Leu Ile Gly Asn Ile Phe Val
            260                 265                 270
Arg Lys Tyr Leu Phe Asn Ser Ala Arg Cys Arg Arg Asp Leu Phe Ala
        275                 280                 285
Thr Gly Gly His Val His Cys His Leu Val Ala Pro Glu Ala Asp Asp
    290                 295                 300
Lys Ala Thr Cys Pro His His Asp Asn Arg His Met Gly Asn Arg Ala
305                 310                 315                 320
Gly Asp His His Ser Leu Ala Ala Leu Leu Arg Pro Gly Ala Arg Arg
                325                 330                 335
Gly Gly Leu Leu Arg Arg Pro Gly Leu Gly Leu Leu Ala Ala Ala Val
            340                 345                 350
Pro Leu Pro Gly Gly Val Ala Pro Gly His Gly Trp Glu Pro Leu Leu
        355                 360                 365
Pro Ala Ser Gln Pro Gly Gly Leu Leu Pro Ala Ala His Val Pro Asp
    370                 375                 380
His Ala Leu Leu Arg Ala His Leu Asp Gln Gly Leu His Glu Val His
385                 390                 395                 400
Pro Arg Arg Val Gln Gly Arg Ala Asn Gly Pro His Ala Ala Glu Glu
                405                 410                 415
Gln Gly Glu Gly His Gln Asp Val Gly Gly Arg Gly His Pro Val Cys
            420                 425                 430
Phe Val Leu Ala Ala Ala Leu Cys His Leu Cys Ala Asp Gln Ile Arg
        435                 440                 445
Leu Gly His Leu Ala Gly Gly Val Arg Asp Pro Gln Glu Gly Asp Ala
    450                 455                 460
Gly Gly Pro Val Ala Gly Leu Leu Glu Gln Leu His Gln Pro His Thr
465                 470                 475                 480
Leu Leu Ser Glu Gln Glu Val Pro Thr Arg Ile Arg Gly His His Gln
                485                 490                 495
Val Ala Gln Leu Leu Arg Thr Pro Gln Val Ser
            500                 505
```

<210> 46
<211> 3688
<212> DNA
<213> Drosophila

<400> 46

```
tgtttgcaga gcccacagca accaatatgt tttcattgtc tgcccggttt gtcgctcaat 60
ttatttgtga tgttgacttt ccaacagtga cgtcagtgga ggccacagac aataccacat 120
catcgcaaca attataaatg gtgtggtggc gaaaaaacgt ccttgtgccc attgttcagt 180
ggggcgaact gaaaagtgga ttttccagcg cgtgccagga aaatgaatga aaatgcggga 240
cattcttgaa gtcgttgcct ttattctgct ttcacacata attatgaaat aaatattagc 300
ttaaaccaaa ttgactacaa cgttatgctc tattaactct tatggtctac ttatgcttaa 360
tgggtttttgg tttaagatgc ttgggcatgt ttatccacca ttacgagtgg ctatatacac 420
ggatgagttg gaatttaaat ggccttaata gctagttagg actggaaaga aatgcatttt 480
aaatttctga taaatggccg cgattcacat acaagtaaaa tcagtctgtt taagagaata 540
```

```
taactaaacc gagctactaa tggattgaaa cctaggttct gagtggtgga caaaaaatct 600
taatagtttt cgaaacactt atttcattag gataccatat acttaaatta atttctccag 660
tttaaagttg attgcaaaaa tgcctaatta aatgtttcag gtaagttttt tggaaatata 720
aaacctttgt accaatatca aattataatg ggttctttag ttaaaattta ataaaaatca 780
aaatcttgaa tagattcaat tacccatttt gacactttag atttttaatt ttagagttaa 840
ttaaatgccg tataaatgtt atcttaacta tagattgcga agtagaaata ggtttgaaat 900
tcgtagcgta gacttggcgt acttacgact agaaacatac atttggcttt acccatgcct 960
ggtataacta tcgaacttat tgtctaagcc tctggtgcct tcacaggatg tccgtgttcc 1020
cctttagaga aactccgcgt cgctaaagga cgtgggacca tgaatcgaga acgcggatct 1080
tttccggctg gtgtgggcca gctatcatac ctccgacatg gctggcacgg taaccacctt 1140
ggtgggacgt tcgcagatta ccacacgggc gttcgtcttg ttggtggacc ttcgggaaac 1200
cgtcaacgaa gtggtcagtg tcctcatcga atcgcaactg ttgtgaagcc gacgaccaac 1260
cgtgtggcag cggttttgct gcgagttgtt gcggtatgac ttgcagcagc atgtgaacca 1320
cttaaaaggc ggaaagcgac tggaaatata ataaaaataa aaatatcaaa gtaagtaaat 1380
taaagcataa gaaaatgtgt aagatttttt aacatgaaat cttaacatac cttaatgtgc 1440
gaaagacctg cgatgagaag aggcaataga ttagtggatt cgccgccgag ttcagcggtg 1500
ccagactttg gatgaaggtg gcgatggcaa tgttggtctg tgagtgtgga atctggccaa 1560
agacctgcag cagatcgaag atgatatacg gcgaccagca gatgatgaac acaaacacga 1620
tggtcaatgt catcttgacc gttttgacct ttgcccgtgg aataatgccc ctcgagctgg 1680
ccctcctggc aggtgcagct ccaaaaccag cacgttctgt taaaatataa ttaacaaata 1740
aaataaaatt atttttggga actccgattt ggcaatatgc atagccagaa gatatcttac 1800
ccgtgggtac aaaaatggaa ccctttgccc aaatcgtctt tacgatgatc gcatagcagg 1860
cagatatgat cagcgcagga atggcaaata gagtggccga caccaggctc atgtacacct 1920
gccaggcgat cggtgaaccc aactcaatcc agcattgcgg atgtccttgg atgagcttct 1980
cctcgtacaa aaccaggatg ggaagcgaaa acaacgccga gatgagccat gcgccagcca 2040
ccaggtgacg ggctcttttc cctaaaagca taaaatttaa atatatattt aagtgaaata 2100
tatttctaat ttaacttaaa tttgtatatt ttttttggtt aattcccgaa ctaaattatc 2160
caagcagtgt aagcagaagc cacaaacagc gctcatcttt ggctaacaac atttggcaat 2220
tagaaaactt ttcatacgac gacttaactt acacgacttt gagaagttca tggggtgtgt 2280
gatggcatcg tatctgtcga tgctcatggc caccagcacg taggtggacg agtatgtgac 2340
gcagacctgc gagaagcgga tggccttgca ggccaggttg cctgcccgcc acgaaatcgt 2400
gatgcgccat atgatgtcgg tgaggacgtt gagcagtccc acgcacagat ctggaaataa 2460
tggactgggc attagaagca agggaagagg tcagtcccaa tcccaaaact cacctgccaa 2520
tgccagctgt ttaatgaagt agttcatccg cgacttgcga ttcttgttga tgaacatcac 2580
gaacagaaca gctgaattgc ccagaacgat gacggtgaac aggatccaga gcacagcaaa 2640
ctgttcggtc tggaagagtg caaagaacaa agtgtggtca catttcatta aacgcctatc 2700
aagttcattg tcaacaaagt gttaagcacg gaaatcacgc ataaaactca atttaatgcc 2760
atccaagcac acacacacac acacgccttg tttatggctc tttaaatagc cataaaaatg 2820
atagagctat gaattcaagc tggttggtgc ggcgggcagg atgtgccggc ggaagtttgc 2880
ttcttccagt tcagtactaa caagactagg acaacacgac atcctgcact caaacacccg 2940
tggacgctca cccacacacc cagacgtgat ggaagacaga cagccaccgt tattccagca 3000
ggaaattgct gccaagatga aggacctgca cttcgaaaaa tatttctttc gacttttaag 3060
gcagtttcac agctcaaaat gcgtggcttg gtatttaatt tatagaactt atactccagg 3120
cattcgaaat tcaattcata gcactcaaat ctaaaattgt aacaattctt atatagaaaa 3180
ctagtagata aagaataatg aaccccttatt acttttactt ttttaactga actgttatta 3240
ttttaaaaat attttaaaca gtgtgaagtg cagatatgac agtgagtagt tctgctagtt 3300
ctagcttatt gccgaagaaa atgaaatatt ctttatacgc ccgcatgcag tttatatagc 3360
agaatatcgc atatatgcta atatgactaa tgatctcttc atggtatcta caaagttttt 3420
tgctgcaaat taaatttgct ctcaaggcat ctccccctat gcccacgtct attgaagtta 3480
agtttgaatt atggcctcgg ggtcttgcag tctgacgcca acaccgatgt tgatgcggac 3540
ttacttaagg gcaggccaag taaacttgtc ccgaaaaaaa agctgaaatc ctgacaggct 3600
tcgtccttgg aactgcgctc agactgaaga cgccccacct tattctacaa ctttttgtcc 3660
ataccaaact tgcagtattc actttttg 3688
```

<210> 47
<211> 1191
<212> DNA
<213> Drosophila

<400> 47

```
atgaaatgtg accacacttt gttctttgca ctcttccaga ccgaacagtt tgctgtgctc 60
tggatcctgt tcaccgtcat cgttctgggc aattcagctg ttctgttcgt gatgttcatc 120
aacaagaatc gcaagtcgcg gatgaactac ttcattaaac agctggcatt ggcagatctg 180
tgcgtgggac tgctcaacgt cctcaccgac atcatatggc gcatcacgat ttcgtggcgg 240
gcaggcaacc tggcctgcaa ggccatccgc ttctcgcagg tctgcgtcac atactcgtcc 300
acctacgtgc tggtggccat gagcatcgac agatacgatg ccatcacaca ccccatgaac 360
ttctcaaagt cgtggaaaag agcccgtcac ctggtggctg gcgcatggct catctcggcg 420
ttgttttcgc ttcccatcct ggttttgtac gaggagaagc tcatccaagg acatccgcaa 480
tgctggattg agttgggttc accgatcgcc tggcaggtgt acatgagcct ggtgtcggcc 540
actctatttg ccattcctgc gctgatcata tctgcctgct atgcgatcat cgtaaagacg 600
atttgggcaa agggttccat ttttgtaccc acggaacgtg ctggttttgg agctgcacct 660
gccaggaggg ccagctcgag ggcattattt ccacgggcaa aggtcaaaac ggtcaagatg 720
acattgacca tcgtgtttgt gttcatcatc tgctggtcgc cgtatatcat cttcgatctg 780
ctgcaggtct ttggccagat tccacactca cagaccaaca ttgccatcgc caccttcatc 840
caaagtctgg caccgctgaa ctcggcggcg aatccactaa tctattgcct cttctcatcg 900
caggtctttc gcacattaag tcgctttccg ccttttaagt ggttcacatg ctgctgcaag 960
tcataccgca acaactcgca gcaaaaccgc tgccacacgg ttggtcgtcg gcttcacaac 1020
agttgcgatt cgatgaggac actgaccact tcgttgacgg tttcccgaag gtccaccaac 1080
aagacgaacg cccgtgtggt aatctgcgaa cgtcccacca aggtggttac cgtgccagcc 1140
atgtcggagc gacgcggagt ttctctaaag gggaacacgg acatcctgtg a          1191

<210> 48
<211> 396
<212> PRT
<213> Drosophila

<400> 48
Met Lys Cys Asp His Thr Leu Phe Phe Ala Leu Phe Gln Thr Glu Gln
 1              5                   10                  15
Phe Ala Val Leu Trp Ile Leu Phe Thr Val Ile Val Leu Gly Asn Ser
            20                  25                  30
Ala Val Leu Phe Val Met Phe Ile Asn Lys Asn Arg Lys Ser Arg Met
        35                  40                  45
Asn Tyr Phe Ile Lys Gln Leu Ala Leu Ala Asp Leu Cys Val Gly Leu
    50                  55                  60
Leu Asn Val Leu Thr Asp Ile Ile Trp Arg Ile Thr Ile Ser Trp Arg
65                  70                  75                  80
Ala Gly Asn Leu Ala Cys Lys Ala Ile Arg Phe Ser Gln Val Cys Val
                85                  90                  95
Thr Tyr Ser Ser Thr Tyr Val Leu Val Ala Met Ser Ile Asp Arg Tyr
            100                 105                 110
Asp Ala Ile Thr His Pro Met Asn Phe Ser Lys Ser Trp Lys Arg Ala
        115                 120                 125
Arg His Leu Val Ala Gly Ala Trp Leu Ile Ser Ala Leu Phe Ser Leu
    130                 135                 140
Pro Ile Leu Val Leu Tyr Glu Glu Lys Leu Ile Gln Gly His Pro Gln
145                 150                 155                 160
Cys Trp Ile Glu Leu Gly Ser Pro Ile Ala Trp Gln Val Tyr Met Ser
                165                 170                 175
Leu Val Ser Ala Thr Leu Phe Ala Ile Pro Ala Leu Ile Ile Ser Ala
            180                 185                 190
Cys Tyr Ala Ile Ile Val Lys Thr Ile Trp Ala Lys Gly Ser Ile Phe
        195                 200                 205
Val Pro Thr Glu Arg Ala Gly Phe Gly Ala Ala Pro Ala Arg Arg Ala
    210                 215                 220
Ser Ser Arg Gly Ile Ile Pro Arg Ala Lys Val Lys Thr Val Lys Met
225                 230                 235                 240
Thr Leu Thr Ile Val Phe Val Phe Ile Ile Cys Trp Ser Pro Tyr Ile
                245                 250                 255
Ile Phe Asp Leu Leu Gln Val Phe Gly Gln Ile Pro His Ser Gln Thr
```

```
            260                 265                 270
Asn Ile Ala Ile Ala Thr Phe Ile Gln Ser Leu Ala Pro Leu Asn Ser
        275                 280                 285
Ala Ala Asn Pro Leu Ile Tyr Cys Leu Phe Ser Ser Gln Val Phe Arg
    290                 295                 300
Thr Leu Ser Arg Phe Pro Pro Phe Lys Trp Phe Thr Cys Cys Cys Lys
305                 310                 315                 320
Ser Tyr Arg Asn Asn Ser Gln Gln Asn Arg Cys His Thr Val Gly Arg
                325                 330                 335
Arg Leu His Asn Ser Cys Asp Ser Met Arg Thr Leu Thr Thr Ser Leu
            340                 345                 350
Thr Val Ser Arg Arg Ser Thr Asn Lys Thr Asn Ala Arg Val Val Ile
        355                 360                 365
Cys Glu Arg Pro Thr Lys Val Val Thr Val Pro Ala Met Ser Glu Arg
    370                 375                 380
Arg Gly Val Ser Leu Lys Gly Asn Thr Asp Ile Leu
385                 390                 395
```

```
<210> 49
<211> 5051
<212> DNA
<213> Drosophila

<400> 49
ttccaatgcg atttaaggta attgactata agtacccatg acgaacttct ggacagttcc 60
ccagtggaaa aaaaaaactg gatttatttc ttatcgcaaa acttgcgata accaatttgt 120
atatgttcat tcagggagcg tgccgatttt ttagtttgcg aaattgagtt gttatgtttt 180
ctgagtcatt gtcgagcggg ttttggtagc cggggattaa atctgctgta tcatcttggt 240
ccagcgttta tttctatttg cttttaattt cacaatcaaa acagcaataa taaaaccgtg 300
ccttatctta cggacatatc tccccaagcg tcgtctcgct ctaatttgga tttttcctct 360
ctctctcgct cttttcttgc agctggcgcc cgacggtggg aatttcatca gctgttgttg 420
ttgttcattt cacagcagca gtcggcctaa taaaaatgca atttttatgg ttaagaaaat 480
agtgcgcgaa cgccgttgtg ttgcacattg gaaattgtgc agaagccaga agaaaatgtg 540
tcttggaaaa tggaaaagca cgagcagcag tagcaccagc agcagctata acattgagcc 600
agtgatgtaa tacagagcaa aagtgaagcg gcggcaattg gcattgcaaa aaagaaacag 660
attatacaat ttgtgttgct gccgcactga actaggacaa cagcagcaac aacagcagca 720
gcaacaacaa ccactagtga taaagaccgg ctggcaacag tgcagcggcc acttgcagca 780
acataatttc aaagtcggag cgtgggaatc gcaaatcacg gcagtggcag caacaacagc 840
aactgcagca acactcgagc actcgtggct ctccggatcg ctcatttcgc tcgccgttct 900
cgcgctggcg tttctccgtt tcgctcgcac ccgcctgtgg agcagtcgcg tatatttgga 960
ggcgttgggg cgaacgctca cgacagttgt gcttggaacg cttcggagtg aagatcgtgc 1020
tcgtccgatc ggtgcgcgtg ataaatacac ggcgaacaaa attgagtcaa gagtcagacg 1080
gtcgcaagac aattgacaag acacgaaaaa tcataataaa cacaagtgga aaagcaactg 1140
cttgagagca aaaaagaaat tgttttttac aaatctcgca cgtaaattag caagcgttcg 1200
caagaaacaa accgcataca atactgaaca caaaaaaaag ggtattcaac aaaaaaccaa 1260
aaatatttgg ctgcaaatgt gaagagacga gagacagcgg atatactacc gagtgaaatt 1320
tgtgtcataa gcagagccaa aacaaaaccg aaaagccaaa ttcaataaaa ggccataagt 1380
gactgtgaaa ctgtaaagtg atgtgtggta atcgaaaaac agcataaaga atggccatag 1440
acctcctgat cctggccctg ctgctcgtct cgttcctgat caacctgctg gcgctgtgtg 1500
ccttttggat aacgcctggc ttgaggacca ccgccaatcg ctttacgatc aacctgctgg 1560
ccatcaatct cattggttgc tgcatcttgg cccccacttt gttcctggga ctgccgggaa 1620
agtctgcgga ggcatccact tccaacgccg agaccctgga gttcttctcg aagccgggaa 1680
accatcaggt gatacttcgt cgcaatggcc agctggtgga acaggacggc gtggtggtgc 1740
gaaggaacat cagcgaaaac ggcgatacgg tggagacctt cttcaagtgc aatgccacct 1800
actgtcgcga gttgaccatc gatgagcgcg gagatggtgg ctttgtcatc acggaaaccg 1860
agacacacga agagaatctc tcggcctttg agagcttacc cacggaggct cccattttgc 1920
cgcccgtcca actgagatgt tggtccatag acatgactgc tgcgctggga gctctggcag 1980
tactgctcgt ggtgggggac acctggtgtg ccgttacgga tcccctgcgc taccacagtc 2040
gcatttcggg cgtgaaaacc tggatcttca tagcactcac ctgggtggtg ggcatactct 2100
```

```
ttggagccct ttccgccttc cgagttttgg actttgaggc ggacgccctg ttcagcagac 2160
aacgccgcct ggcagtcacg tacttcaaca tcagcagcac caatagtatt ttcggagtgg 2220
tgtacgccag tgtctacttc atcgtgatca tcctcctgcc cttcggcttc gtgtgcggaa 2280
tgtactggag gatattcagt gaggcgcgcg gaaatgggct gcgcatgcgc cagaatggct 2340
cttcgccgct gctgcagagc gccttgaatc tgacggctgg ccagcaggcg gctcaagcca 2400
accagttctc caatagcttg tgtgttcatc ggcattccat atcctcggcc agttcgcatg 2460
gcggcaacag cagcttgggc ttgggcggtc tgcaaatgca aattgatcag cgccaacagc 2520
cgaggagttc gcccagttgt ctccgtcggg actcggccgc gaaagtactt ttgcccacca 2580
tctcggatga tggtggctca gatgcggaaa gtggagccgg agttcaactg atgcccgtac 2640
aggagcactc actgtcagat agaaatcaga acattatgct tactctgcaa acagccagtg 2700
gtgaaataaa gaggaactac tccgccagac agctgcccct gttgggcact tcttcgcagg 2760
atctgaggga gactaatcgt ctgcagggca tacgccaggt gcacagttcg cccaatctcc 2820
acaagtacac ggaactgcgc caggactcgc tcagtgagga gtgtggttct ccgcatcttt 2880
tgggtcacgc ccagcggcag cagcagcagc aattgcacct gcaacatcag cagcaacagc 2940
atcatcacca ccagcaacat catccgcact tcagttcgcc gcgccaccag caacacggcc 3000
acgccctcca gatacccgcc attcatgcct cccccaaggc tctgagttac atgagctcct 3060
tgcgacaccg tttgagcaac gccagttcgc tattcaaata ccgtgaggag tcgagagcag 3120
cccgcatcag catcctggtg gtggtgatgt ttgtggtctc gtacctgccc ttcggactgc 3180
tggtactcct gcaatcgcga ctgtcggctg ctaactttgg cggatcctcc cagctggcca 3240
tcttcatgat cctgctggcc aatctcagtt cgccattcat cttcgcctac aggaacaaga 3300
gggtgaggcg aggagtcaag cggctgtttg gtctggattc gtcgtcgggc ctgcagcgca 3360
actgcagcag cagtgtaaag accaatggca ccgcaggtcc tgctgcaagt ggtgcacaat 3420
tgcagcgcaa cagcagcaag ttgtcgcagt acagcagcaa cagctgcaag tatctcaccc 3480
cccagagttc cctggtcagt caagtgccgg tgcatacgac cttgacgtta aggcccaaca 3540
gcagctgcag caccatcatc aatttcggtg gcgccagggg ctcagcggac tccgatgagc 3600
agccaccggc aacaccacca cctactgtgg cagttgctcc accgcccaca cgaccgcagc 3660
gccccaagct gcagaggggc atcaccatcg ttgagcacat agctataaca ccaaccatgc 3720
cccaaaagtt ccagaatcgg cgcgccaggc tcttcgacat gttcttccgc agctccaaga 3780
aactgcaggc gggctgccag agccagagtt tgcccactga ggtctaaacg gggggttatg 3840
ataactatac atataggaag gtaaacaaac actaaacttg cttgctgaag aattctaaga 3900
acagtattcg taatttgtaa gtctagcaat taacaaagac aactgtttga atttgagttt 3960
taagaccaga tcatgctgtc ctagaagctt ctcatgccaa aaaacaaaac cacaaagaaa 4020
gtcaaactag aaagtttgag atgctttaaa ttaacaaaaa aagaaaagaa aaaaacaaaa 4080
aacaaaacaa agtacctttt aaaaaataga aaactgtgta caagtcaaat taacaaagat 4140
tgagagtgat gggagttata atcttagggt ggatatcaaa ctgaaaagtg gaaccaagcc 4200
aagccaagtt gagatcgtaa tcaatgtacc taagcaactt gcatatgaac atttaatgtt 4260
tttactcaga tgaaatcaca ggaaaatagc aaagacagca cagcttattc gaacttcccc 4320
aactagttta ttatttattg aaacaattgt aaatatttat tagacatcca cgccgcaaag 4380
ctttaactcc aaaaacattc tggctaaaca agaagagacg catacattcg attcataagt 4440
taattaagtc aatgtgcaaa gcatatgtgt attctagagt ttaagttgta gagccgagta 4500
gtcttacacc gagcacctag tgccgagcag tgctttttca cctgcccaac tggagcgtct 4560
gttgcaggcg gcattttcat attgaaaatg cattttggcc agtgcacagt tggttttcca 4620
gttggccaga atgtcagtca gcagttagtt agcctttggt tagccaacgg ctattgggcc 4680
attaagtcaa tttgcttttt agaacaccgc gggcattgcg tgcagctaaa tacctcatta 4740
gttataattt gaagtgcgag atggatggcg gcaataggtc aaaggtcgga gccatcggca 4800
cggcccatgc ataatgcatg gaaaatgttc gattgcccct ggcatttgta tcatttgctt 4860
tactatgctg caggccagaa agctttatac cttaagttct cctgctttga tttcccttgg 4920
ctgcacaaaa aggcaacaaa agcaacacac actcacacaa cacaaaaaaa cctcgaattt 4980
atatttctaa agtttatatt ttataaagag tgtgtgcata ggaaatatct ataaatattc 5040
atgcatgtgt g                                                      5051
```

<210> 50
<211> 3051
<212> DNA
<213> Drosophila

<400> 50

```
cttcggagtg aagatcgtgc tcgtccgatc ggtgcgcgtg ataaatacac ggcgaacaaa 60
```

```
attgagtcaa gagtcagacg gtcgcaagac aattgacaag acacgaaaaa tcataataaa 120
cacaagtgga aaagcaactg cttgagagca aaaaagaaat tgttttttac aaatctcgca 180
cgtaaattag caagcgttcg caagaaacaa accgcataca atactgaaca caaaaaaaag 240
ggtattcaac aaaaaaccaa aaatatttgg ctgcaaatgt gaagagacga gagacagcgg 300
atatactacc gagtgaaatt tgtgtcataa gcagagccaa aacaaaaccg aaaagccaaa 360
ttcaataaaa ggccataagt gactgtgaaa ctgtaaagtg atgtgtggta atcgaaaaac 420
agcataaaga atggccatag acctcctgat cctggccctg ctgctcgtct cgttcctgat 480
caacctgctg gcgctgtgtg ccttttggat aacgcctggc ttgaggacca ccgccaatcg 540
ctttacgatc aacctgctgg ccatcaatct cattggttgc tgcatcttgg cccccacttt 600
gttcctggga ctgccgggaa agtctgcgga ggcatccact tccaacgccg agaccctgga 660
gttcttctcg aagccgggaa accatcaggt gatacttcgt cgcaatggcc agctggtgga 720
acaggacggc gtggtggtgc gaaggaacat cagcgaaaac ggcgatacgg tggagacctt 780
cttcaagtgc aatgccacct actgtcgcga gttgaccatc gatgagcgcg gagatggtgg 840
ctttgtcatc acggaaaccg agacacacga agagaatctc tcggcctttg agagcttacc 900
cacggaggct cccattttgc cgcccgtcca actgagatgt tggtccatag acatgactgc 960
tgcgctggga gctctggcag tactgctcgt ggtgggggac acctggtgtg ccgttacgga 1020
tcccctgcgc taccacagtc gcatttcggg cgtgaaaacc tggatcttca tagcactcac 1080
ctgggtggtg ggcatactct ttggagccct ttccgccttc cgagttttgg actttgaggc 1140
ggacgccctg ttcagcagac aacgccgcct ggcagtcacg tacttcaaca tcagcagcac 1200
caatagtatt ttcggagtgg tgtacgccag tgtctacttc atcgtgatca tcctcctgcc 1260
cttcggcttc gtgtgcggaa tgtactggag gatattcagt gaggcgcgcg gaaatgggct 1320
gcgcatgcgc cagaatggct cttcgccgct gctgcagagc gccttgaatc tgacggctgg 1380
ccagcaggcg gctcaagcca accagttctc caatagcttg tgtgttcatc ggcattccat 1440
atcctcggcc agttcgcatg gcggcaacag cagcttgggc ttgggcggtc tgcaaatgca 1500
aattgatcag cgccaacagc cgaggagttc gcccagttgt ctccgtcggg actcggccgc 1560
gaaagtactt ttgcccacca tctcggatga tggtggctca gatgcggaaa gtggagccgg 1620
agttcaactg atgcccgtac aggagcactc actgtcagat agaaatcaga acattatgct 1680
tactctgcaa acagccagtg gtgaaataaa gaggaactac tccgccagac agctgcccct 1740
gttgggcact tcttcgcagg atctgaggga gactaatcgt ctgcagggca tacgccaggt 1800
gcacagttcg cccaatctcc acaagtacac ggaactgcgc caggactcgc tcagtgagga 1860
gtgtggttct ccgcatcttt tgggtcacgc ccagcggcag cagcagcagc aattgcacct 1920
gcaacatcag cagcaacagc atcatcacca ccagcaacat catccgcact tcagttcgcc 1980
gcgccaccag caacacggcc acgccctcca gatacccgcc attcatgcct cccccaaggc 2040
tctgagttac atgagctcct tgcgacaccg tttgagcaac gccagttcgc tattcaaata 2100
ccgtgaggag tcgagagcag cccgcatcag catcctggtg gtggtgatgt ttgtggtctc 2160
gtacctgccc ttcggactgc tggtactcct gcaatcgcga ctgtcggctg ctaactttgg 2220
cggatcctcc cagctggcca tcttcatgat cctgctggcc aatctcagtt cgccattcat 2280
cttcgcctac aggaacaaga gggtgaggcg aggagtcaag cggctgtttg gtctggattc 2340
gtcgtcgggc ctgcagcgca actgcagcag cagtgtaaag accaatggca ccgcaggtcc 2400
tgctgcaagt ggtgcacaat tgcagcgcaa cagcagcaag ttgtcgcagt acagcagcaa 2460
cagctgcaag tatctcaccc cccagagttc cctggtcagt caagtgccgg tgcatacgac 2520
cttgacgtta aggcccaaca gcagctgcag caccatcatc aatttcggtg gcgccagggg 2580
ctcagcggac tccgatgagc agccaccggc aacaccacca cctactgtgg cagttgctcc 2640
accgcccaca cgaccgcagc gccccaagct gcagaggggc atcaccatcg ttgagcacat 2700
agctataaca ccaaccatgc cccaaaagtt ccagaatcgg cgcgccaggc tcttcgacat 2760
gttcttccgc agctccaaga aactgcaggc gggctgccag agccagagtt tgcccactga 2820
ggtctaaacg gggggttatg ataactatac atataggaag gtaaacaaac actaaacttg 2880
cttgctgaag aattctaaga acagtattcg taatttgtaa gtctagcaat taacaaagac 2940
aactgtttga atttgagttt taagaccaga tcatgctgtc ctagaagctt ctcatgccaa 3000
aaaacaaaac cacaaagaaa gtcaaactag aaagtttgag atgctttaaa t          3051
```

<210> 51
<211> 798
<212> PRT
<213> Drosophila

<400> 51

```
Met Ala Ile Asp Leu Leu Ile Leu Ala Leu Leu Leu Val Ser Phe Leu
 1               5                  10                  15
Ile Asn Leu Leu Ala Leu Cys Ala Phe Trp Ile Thr Pro Gly Leu Arg
```

```
            20                      25                      30
Thr Thr Ala Asn Arg Phe Thr Ile Asn Leu Leu Ala Ile Asn Leu Ile
        35                      40                      45
Gly Cys Cys Ile Leu Ala Pro Thr Leu Phe Leu Gly Leu Pro Gly Lys
    50                      55                      60
Ser Ala Glu Ala Ser Thr Ser Asn Ala Glu Thr Leu Glu Phe Phe Ser
65                      70                      75                      80
Lys Pro Gly Asn His Gln Val Ile Leu Arg Arg Asn Gly Gln Leu Val
                85                      90                      95
Glu Gln Asp Gly Val Val Val Arg Arg Asn Ile Ser Glu Asn Gly Asp
            100                     105                     110
Thr Val Glu Thr Phe Phe Lys Cys Asn Ala Thr Tyr Cys Arg Glu Leu
        115                     120                     125
Thr Ile Asp Glu Arg Gly Asp Gly Gly Phe Val Ile Thr Glu Thr Glu
    130                     135                     140
Thr His Glu Glu Asn Leu Ser Ala Phe Glu Ser Leu Pro Thr Glu Ala
145                     150                     155                     160
Pro Ile Leu Pro Pro Val Gln Leu Arg Cys Trp Ser Ile Asp Met Thr
                165                     170                     175
Ala Ala Leu Gly Ala Leu Ala Val Leu Leu Val Val Gly Asp Thr Trp
            180                     185                     190
Cys Ala Val Thr Asp Pro Leu Arg Tyr His Ser Arg Ile Ser Gly Val
        195                     200                     205
Lys Thr Trp Ile Phe Ile Ala Leu Thr Trp Val Val Gly Ile Leu Phe
    210                     215                     220
Gly Ala Leu Ser Ala Phe Arg Val Leu Asp Phe Glu Ala Asp Ala Leu
225                     230                     235                     240
Phe Ser Arg Gln Arg Arg Leu Ala Val Thr Tyr Phe Asn Ile Ser Ser
                245                     250                     255
Thr Asn Ser Ile Phe Gly Val Val Tyr Ala Ser Val Tyr Phe Ile Val
            260                     265                     270
Ile Ile Leu Leu Pro Phe Gly Phe Val Cys Gly Met Tyr Trp Arg Ile
        275                     280                     285
Phe Ser Glu Ala Arg Gly Asn Gly Leu Arg Met Arg Gln Asn Gly Ser
    290                     295                     300
Ser Pro Leu Leu Gln Ser Ala Leu Asn Leu Thr Ala Gly Gln Gln Ala
305                     310                     315                     320
Ala Gln Ala Asn Gln Phe Ser Asn Ser Leu Cys Val His Arg His Ser
                325                     330                     335
Ile Ser Ser Ala Ser Ser His Gly Gly Asn Ser Ser Leu Gly Leu Gly
            340                     345                     350
Gly Leu Gln Met Gln Ile Asp Gln Arg Gln Gln Pro Arg Ser Ser Pro
        355                     360                     365
Ser Cys Leu Arg Arg Asp Ser Ala Ala Lys Val Leu Leu Pro Thr Ile
    370                     375                     380
Ser Asp Asp Gly Gly Ser Asp Ala Glu Ser Gly Ala Gly Val Gln Leu
385                     390                     395                     400
Met Pro Val Gln Glu His Ser Leu Ser Asp Arg Asn Gln Asn Ile Met
                405                     410                     415
Leu Thr Leu Gln Thr Ala Ser Gly Glu Ile Lys Arg Asn Tyr Ser Ala
            420                     425                     430
Arg Gln Leu Pro Leu Leu Gly Thr Ser Ser Gln Asp Leu Arg Glu Thr
        435                     440                     445
Asn Arg Leu Gln Gly Ile Arg Gln Val His Ser Ser Pro Asn Leu His
    450                     455                     460
Lys Tyr Thr Glu Leu Arg Gln Asp Ser Leu Ser Glu Glu Cys Gly Ser
465                     470                     475                     480
Pro His Leu Leu Gly His Ala Gln Arg Gln Gln Gln Gln Gln Leu His
                485                     490                     495
Leu Gln His Gln Gln Gln Gln His His His His Gln Gln His His Pro
```

```
            500                 505                 510
His Phe Ser Ser Pro Arg His Gln Gln His Gly His Ala Leu Gln Ile
        515                 520                 525
Pro Ala Ile His Ala Ser Pro Lys Ala Leu Ser Tyr Met Ser Ser Leu
    530                 535                 540
Arg His Arg Leu Ser Asn Ala Ser Ser Leu Phe Lys Tyr Arg Glu Glu
545                 550                 555                 560
Ser Arg Ala Ala Arg Ile Ser Ile Leu Val Val Val Met Phe Val Val
                565                 570                 575
Ser Tyr Leu Pro Phe Gly Leu Leu Val Leu Leu Gln Ser Arg Leu Ser
            580                 585                 590
Ala Ala Asn Phe Gly Gly Ser Ser Gln Leu Ala Ile Phe Met Ile Leu
        595                 600                 605
Leu Ala Asn Leu Ser Ser Pro Phe Ile Phe Ala Tyr Arg Asn Lys Arg
    610                 615                 620
Val Arg Arg Gly Val Lys Arg Leu Phe Gly Leu Asp Ser Ser Ser Gly
625                 630                 635                 640
Leu Gln Arg Asn Cys Ser Ser Ser Val Lys Thr Asn Gly Thr Ala Gly
                645                 650                 655
Pro Ala Ala Ser Gly Ala Gln Leu Gln Arg Asn Ser Ser Lys Leu Ser
            660                 665                 670
Gln Tyr Ser Ser Asn Ser Cys Lys Tyr Leu Thr Pro Gln Ser Ser Leu
        675                 680                 685
Val Ser Gln Val Pro Val His Thr Thr Leu Thr Leu Arg Pro Asn Ser
    690                 695                 700
Ser Cys Ser Thr Ile Ile Asn Phe Gly Gly Ala Arg Gly Ser Ala Asp
705                 710                 715                 720
Ser Asp Glu Gln Pro Pro Ala Thr Pro Pro Pro Thr Val Ala Val Ala
                725                 730                 735
Pro Pro Pro Thr Arg Pro Gln Arg Pro Lys Leu Gln Arg Gly Ile Thr
            740                 745                 750
Ile Val Glu His Ile Ala Ile Thr Pro Thr Met Pro Gln Lys Phe Gln
        755                 760                 765
Asn Arg Arg Ala Arg Leu Phe Asp Met Phe Phe Arg Ser Ser Lys Lys
    770                 775                 780
Leu Gln Ala Gly Cys Gln Ser Gln Ser Leu Pro Thr Glu Val
785                 790                 795
```

```
<210> 52
<211> 3919
<212> DNA
<213> Drosophila

<400> 52
attctattgt cgtaggcgtg gcaaaaggtt ttttggtaaa tcgatagaaa tgtacaagag 60
aatgtaatta atatgaaaag gtaacaaaaa attttgcaaa agtgtgggca tggcagtttt 120
ggtcggtttg ttggcgtaag ggcattatcg gtattttcct attttcagct atatgatgca 180
gtcgcctaat tttgatcaaa ctataactaa aattgtatat tgtcagaaac aacgagcatg 240
aatttatagt caatttatct ttttattatt tcttatctaa agttatatgc atatagaata 300
aaaacagata cgagcttcct tagaagcgga cggaaataat gtctcgagat ctggaccttc 360
attaagacag acgaacatta ggacaaggcc aggtctactc agctagggat cctgatcaat 420
gtatatgtat ttactttaaa ggggcgtaaa cttgtatttt acatgttaca tactttcaag 480
cgaatctagt aaaccctttt aatttacaag taacgaatat aaaaagttga atggtagcgg 540
gagaaaatta attgtataat ataatattat ctttgttaat ccagagtctt ggcaacttgt 600
gcacacaatt tttttttttt ttttggaaac tgaatcagtt aattgtgcca aaatttagag 660
tgtagatttc ttggaaatct caatgaattt tatggtcaag ttccaagaag tgaaatgctt 720
aacaaacaac aaaaacaaaa cacacaaaag caaacaaatt atgcagaatt atcggatgta 780
ttttatatgt atgtatgtac atacatatac aaagaataat aaatgggtgg gtctgttggc 840
ttctatggaa tttccaattt aagctgcaat acatgtacat tttaacgtct ttctgtacgc 900
```

tctcaaataa aaaaacacaa ttttaggatt aataaaaaa cagatatgta cgtcttctat 960
tcccatatga agaaaaatat tcatatgatt ttttgtttgg caggaaatgc gaatagttat 1020
tggatcgttc accgcatttc ttttgctgtt attgcaaaac tcaaatgccg aaattcccgg 1080
ttgcgacttc ttcgacaccg tagatatttc aaaagcgcca agattctcga acggatcgta 1140
cctctacgaa ggcttgctga tccccgccca tttgacagct gaatatgact acaagctcct 1200
ggccgacgat tcgaaggaga aggtggcgag ccacgtacga ggatgtgcct gccacctcag 1260
gccatgcatt cggttttgtt gcccccagta ccaaaagatg caaaagagca agtgctacgg 1320
cgacatgtcg gaggacgagc tgaacaagca cgatcccttc gtgaacgtga cgctcagcga 1380
cgggtcggtg gtcaggagac acttcaagga ggatctgatc gtgcagtcgg atctggccaa 1440
gcccggatgt ccccggatgt actttctgaa tcacgaactg ccgggcaatg aattcactct 1500
gtttgaggtt actttcctgt gtttctgtca cttctgaaga tcgtaacctt ttttgaatct 1560
ctgtagaacg gctctcttct gcgtcactgg gacaaagtgg aactgagcaa gccgggagta 1620
ctgcgtccag catctttcat ttaaggatga tagtatccga attgcacccc acttctgtcc 1680
actttcgtct gagcactcca gaacgtggaa gaccgttggt aagctgaatt ccctgcttga 1740
ctttaaggtt tttagttaat gtccattttc acagcgatag tgatatcctt gatatgcata 1800
atcctaacga tcagcgtgta cctctatgtt gagaagcttc gcaacctgca cggaaagtgt 1860
ttcatctgct acttggcctc cttgttcctg ggatatttct tcctggtcct caacgtgtgg 1920
aaatactcct ccggcttctg cgttacagca ggtttgcttg tggcgattac aagataaagg 1980
acgaccaatt tcacatttcg atatcatgat cagggttcct gggctacttc tccgtcatcg 2040
ccgcgttctt ttggctttct gtcatcagtc ttaccctctg gaattcattt agcggcaact 2100
ccagctggtt gaaccgtttc ttgccacaga atcggttcct ttcttacaat ttatacgcct 2160
ggggcatggc gctgctcctg actgcaatca cctatatagc tgaccaggtg gttaagaacg 2220
agaagttgag accccgcgtg ggcgtcggca aaaattgttg gatctacagt aagttactat 2280
tcctataata acaattgcgg tgtaactcgt tttttttta accaacagct ggggatatga 2340
ctgtcatgat ctacttctat ggcccgatgc tactgataat tgttttcaac ataacaatgt 2400
ttgtcctgac ggcttttcgt ataatgaaag tgaagaagga agcacaaaac tttactcaac 2460
agcaaaaaac aaccaatagg ctcaattcgg acaagcagac gtaataattt taccatattt 2520
tatacagttc taaatccaaa agtttaaatt ttaaattatt ttctatggct tttaaagtta 2580
cgccctattc ctgcgactct ttatcatcat gggcttgtcc tggagcttgg aaataatctc 2640
gtttttgttg agcaaaaatc aggcttgggc caaggctttt atggtggctg attacttcaa 2700
ttggtcacag ggcaccgtca tatttctgct atttgttctg aggcccagca cactgaagct 2760
actgaaggaa cggtaagtta tcggttcttc cataggaggt ccaattaacc tggttctatc 2820
atctggctga gcaggattaa gggtgggagg gatgaagccg gcgccagcga tgaacatatc 2880
tcacttcaga atacgaaaat tgaccccagt gtttttaat atatgggtaa tgtgaaaata 2940
cagtttatta gttcatttgg tgttggaaac cttctatttt gaatcactaa caggcagcct 3000
aatgatcgtg accaaaggct gctttgcttg ctcattattg ctcacatatc gtatgctcat 3060
atgggctacg aatatttgct acttatcact tttcaagtat cttgatttct gcacggtgtc 3120
taagcatatc taagttttaa ttaaatgttg aatgtaattg tcctttaagt tataactttt 3180
tgcaataata atctattcca atccatccta gacgtgccat caaagcgaga tcaaagcaac 3240
tactacatgg ggatatgtac gaatcttata ggcatgaata ctcggcattt atataccctg 3300
gtctattatt atggatggag gataaaaact caggaaaata cttttcgcat cagcacttat 3360
gtatgctcac tcgtgataat aatttcgtaa taattcgtac tgtttttact gcctcaactt 3420
aattcaaaga actacaaatg attatattat gactttttat gacaatgact aactaaaaat 3480
aaaaatttct gaatatttat aaacaacttt aaaatgttat taacttagga taataataat 3540
aacacagttt taatttatca agcacatcgt taaaaaacaa taaattactt gaatataaat 3600
ttaaaaaaaa aatggaatat tggtttacat taacggtatg caaaaaatca taaaaagtca 3660
tagaactgaa cgatagagaa aattacaaaa aaacgatcta gtcgagttcc tcgactataa 3720
gatactcatt tctcagctag ttctttcatt taaaaatttc tctttcgttt cgatagtaat 3780
tggcaagaaa aaataataaa atacaacaaa ttaatgcatt tttgaaaatt atgggcgtga 3840
ctgatttggt cggttagtgg acgtgtcaat attatgaagc aaacgtggac actacaaact 3900
gaaatatata tatatatat 3919

<210> 53
<211> 1547
<212> DNA
<213> Drosophila

<400> 53
caggaaatgc gaatagttat tggatcgttc accgcatttc ttttgctgtt attgcaaaac 60
tcaaatgccg aaattcccgg ttgcgacttc ttcgacaccg tagatatttc aaaagcgcca 120

```
agattctcga acggatcgta cctctacgaa ggcttgctga tccccgccca tttgacagct 180
gaatatgact acaagctcct ggccgacgat tcgaaggaga aggtggcgag ccacgtacga 240
ggatgtgcct gccacctcag gccatgcatt cggttttgtt gcccccagta ccaaaagatg 300
caaaagagca agtgctacgg cgacatgtcg gaggacgagc tgaacaagca cgatcccttc 360
gtgaacgtga cgctcagcga cgggtcggtg gtcaggagac acttcaagga ggatctgatc 420
gtgcagtcgg atctggccaa gcccggatgt ccccggatgt actttctgaa tcacgaactg 480
ccgggcaatg aattcactct gtttgagaac ggctctcttc tgcgtcactg ggacaaagtg 540
gaactgagca agccgggagt actgcgtcca gcatctttca tttaaggatg atagtatccg 600
aattgcaccc cacttctgtc cactttcgtc tgagcactcc agaattaatg tccattttca 660
cagcgatagt gatatccttg atatgcataa tcctaacgat cagcgtgtac ctctatgttg 720
agaagcttcg caacctgcac ggaaagtgtt tcatctgcta cttggcctcc ttgttcctgg 780
gatatttctt cctggtcctc aacgtgtgga aatactcctc cggcttctgc gttacagcag 840
ggttcctggg ctacttctcc gtcatcgccg cgttcttttg gctttctgtc atcagtctta 900
ccctctggaa ttcatttagc ggcaactcca gctggttgaa ccgtttcttg ccacagaatc 960
ggttcctttc ttacaattta tacgcctggg gcatggcgct gctcctgact gcaatcacct 1020
atatagctga ccaggtggtt aagaacgaga agttgagacc ccgcgtgggc gtcggcaaaa 1080
attgttggat ctacactggg gatatgactg tcatgatcta cttctatggc ccgatgctac 1140
tgataattgt tttcaacata acaatgtttg tcctgacggc ttttcgtata atgaaagtga 1200
agaaggaagc acaaaacttt actcaacagc aaaaaacaac caataggctc aattcggaca 1260
agcagactta cgccctattc ctgcgactct ttatcatcat gggcttgtcc tggagcttgg 1320
aaataatctc gtttttgttg agcaaaaatc aggcttgggc caaggctttt atggtggctg 1380
attacttcaa ttggtcacag ggcaccgtca tatttctgct atttgttctg aggcccagca 1440
cactgaagct actgaaggaa cggattaagg gtgggaggga tgaagccggc gccagcgatg 1500
aacatatctc acttcagaat acgaaaattg accccagtgt tttttaa               1547
```

<210> 54
<211> 192
<212> PRT
<213> Drosophila

<400> 54

```
Met Arg Ile Val Ile Gly Ser Phe Thr Ala Phe Leu Leu Leu Leu Leu
 1               5                  10                  15
Gln Asn Ser Asn Ala Glu Ile Pro Gly Cys Asp Phe Phe Asp Thr Val
            20                  25                  30
Asp Ile Ser Lys Ala Pro Arg Phe Ser Asn Gly Ser Tyr Leu Tyr Glu
        35                  40                  45
Gly Leu Leu Ile Pro Ala His Leu Thr Ala Glu Tyr Asp Tyr Lys Leu
    50                  55                  60
Leu Ala Asp Asp Ser Lys Glu Lys Val Ala Ser His Val Arg Gly Cys
65                  70                  75                  80
Ala Cys His Leu Arg Pro Cys Ile Arg Phe Cys Cys Pro Gln Tyr Gln
                85                  90                  95
Lys Met Gln Lys Ser Lys Cys Tyr Gly Asp Met Ser Glu Asp Glu Leu
            100                 105                 110
Asn Lys His Asp Pro Phe Val Asn Val Thr Leu Ser Asp Gly Ser Val
        115                 120                 125
Val Arg Arg His Phe Lys Glu Asp Leu Ile Val Gln Ser Asp Leu Ala
    130                 135                 140
Lys Pro Gly Cys Pro Arg Met Tyr Phe Leu Asn His Glu Leu Pro Gly
145                 150                 155                 160
Asn Glu Phe Thr Leu Phe Glu Asn Gly Ser Leu Leu Arg His Trp Asp
                165                 170                 175
Lys Val Glu Leu Ser Lys Pro Gly Val Leu Arg Pro Ala Ser Phe Ile
            180                 185                 190
```

<210> 55
<211> 3640
<212> DNA

<213> Drosophila

<400> 55
aattaatgca tttttgaaaa ttatgggcgt gactgatttg gtcggttagt ggacgtgtca 60
atattatgaa gcaaacgtgg acactacaaa ctgaaatata tatatatata tcatttcatt 120
gcccttaagt ctgaaaaacg gctttaaaat tgagaagtta ctttaagaga tattcgaggt 180
attcggtgac agcagcgcac attatcgact ggccaatggg aatattagat ttgtagaata 240
gggtaatatg taaacgtaaa cctttcgatt gtataaagca ttcaattcct aatatggata 300
tgtccgttca ccagcctgta ttaactttgc gtgttattac ttaaatttat ctttctatgt 360
acatgccaga aactgataga atggcacggt taccgccgat agtcgctgag tatggggtat 420
cttatagtcg agcacattct cgttctccct tgacatcgca tatatcggtt gagttgttcg 480
cattttggga tcagtcatct taaaactctg gtcgcataaa catgtttttt gtgttagtcg 540
ttccatttga cccttgcgta agtatttata ttcaaagttt ttacgatttt ggatcgttaa 600
gcataagaag gaccactcga gtacgatttg accgttttta tacaatttcg tcctcttaga 660
cgaataaata ttgtgcctgc cttgtttcta aaaatgaact tgcttcagct tcgtagagaa 720
tgcttagtca taacagatgg atttgttgaa tttattttta acagttgata agttatactt 780
ttataaaaaa aaaaaaacaa cttactcaga atgctatagc acatgactca atttttgtt 840
tttcataagt atttagatca gttttcagtt ttgcggaggt tattttgcta ctttatttct 900
aattgtgaac acgtaacttg gtttaaaacc gtatatactt acagaatagc atacatcatt 960
tgaattcagt tgcctatcat attaatttcc tttgcagaaa atgcggatat tgttaattgc 1020
agttcttttt ctgttgatgc caaaatcaaa tgccgaaatt cccggttgcg acttcttcga 1080
caccgtagat atttcaaaag cgccaagatt ctcgaacgga tcgtacctct acgaaggctt 1140
gctgatcccc gcccatttga cagctgaata tgactacaag ctcctggccg acgattcgaa 1200
ggagaaggtg gcgagccacg tacgaggatg tgcctgccac ctcaggccat gcatccggtt 1260
ttgttgcccc cagtaccaaa agatgcaaaa gagcaagtgc tatggcgaca tgtcggagga 1320
cgagctgaac aagcacgatc ccttcgtgaa cgtgacgctc agcgacgggt cggtggtcag 1380
gagacacttc aaggaggatc tgatcgtgca gtcggatctg gccaagcccg gatgtccccg 1440
gatgtacttt ttgaatcacg aactgccggg caatgaattc actctgtttg aggttacttt 1500
cctgtgtttc tgtcacttct gaagatcgta accttttttg aatctctgta gaacggctct 1560
cttctgcgtc actgggacaa agtggaactg agcaagcggg agtactgcgt ccagcatctt 1620
tcatttaagg atgatagtat ccgaattgca ccccacttct gtccactttc gtctgagcac 1680
tccagaacgt ggaagaccgt tggtaagctg aattccctgc ttgactttaa ggtttttagt 1740
taatgtcctt tctcacagcg atagtgatat ccttgatatg cataatccta acgatcagcg 1800
tgtacctcta tgtcgagaag cttcgcaacc tgcacggaaa gtgtttcatc tgctacttgg 1860
cctctttgtt cttgggctat ttcttcctgg tcctcaacgt gtggaaatac tcctccggct 1920
tctgcgttac agcaggtttg cttgtggcga ttacaagata aaggacgact aatttcacat 1980
ttcgatatca tgatcagggt tcctgggcta cttctccgtc atggccgcgt tcttttggct 2040
ttctgtgatt ggaatccact tgcggattaa gtttagtcta gcctctaact gtttgcaccg 2100
attgctgcca gaaaatccgt tccgtgcata caacttatac gcctggggta taccgctaat 2160
catgactgca atcacctata cagctgacca ggtggttaag aacgagaagt tgagaccccg 2220
cgtgggcgtc ggcaaaaatt gttggatcta cagtaagtta ctattcctat aattacaatt 2280
acggtgtaac tcgttttgtt aaccaacagc tggcgatatg actgtcatga tctacttcta 2340
tggcccgatg ctgttgctaa tcgctttcaa cataataatg tttgtccttt cggcgattta 2400
catatataac ataaagaaaa atgtgaaagg ccttgtccac aagcaacaaa caaaccaaca 2460
gataaatgac caacaaatgt aattatatta tattataaga taataactat aatgtcaaat 2520
ggcacttttt gtaatctttt ataggtttgc tatattcctg cgacttttca tcttaatggg 2580
tttgtcgtgg agctttgaga tattatcctt tttgttaacc aaacagcaag cttgggctag 2640
ggctttaatg gtggctgact actttaattg gtcccagggt accatcatat tcgtgctttt 2700
tattttaaag cctagcattc taaaacttat aatagcaggg taagtaaagt attacaaata 2760
accttgaaat ttgaccgtct catacgtatg aacaatcttt tgatttaggg gacgtcaaaa 2820
cctcccagga agtcaccata attcgagatc aaaagcagct cgatataatt caactcatac 2880
ggcttgtgaa ggatcaattg cggatccaaa cgcttactgc taaaaaaatg gaattcaata 2940
agatgcaacg tctctgttat tgatcacatt attttaattt tctatggagc aaaagtttaa 3000
aaaaaatatt tgttatcaag acaataaata atgtcttgga acatataaag caaaatatat 3060
taaagggtac aaaaaacagg aaaagtatgg ggcaatacag caaaatattt ttcggatcat 3120
caactttttt gtaagacaaa atgtagaaat atttctggca ccgttgtttc aaaccagggt 3180
ttgagtcaac aatgctggcc caccactggc aagtaccaaa cgattacaga aaactttttc 3240
taatcgatta agcatcgggg gtcatcgatg tttctccacc actacaattt catgttgttt 3300
ttttagtcgc ggctgtttga cgtttgtcgc tttcctccgt tgcgaatata atatattacg 3360
tagctcattt ttatacaaac ggaattacga gcgcaacgac gacagcaaca ctagtagcac 3420

taatcgtaag cgcaggggcc aaaaattaaa ttgcgtttgc ggccgcaaag atttgatgac 3480
gtcgcatacg ccgtcttcta gggcgtaaaa agcaaagcaa agcaaacaaa cgcgaaagcg 3540
aaacgtgtaa acggcgtaga agcgataaac gcgactcaaa tacgcagcag ataaaataca 3600
atacgcgaga agagaaaagt cacgggaaat attgttcata 3640

<210> 56
<211> 1089
<212> DNA
<213> Drosophila

<400> 56
atgcggatat tgttaattgc agttcttttt ctgttgatgc caaaatcaaa tgccgaaatt 60
cccggttgcg acttcttcga caccgtagat atttcaaaag cgccaagatt ctcgaacgga 120
tcgtacctct acgaaggctt gctgatcccc gcccatttga cagctgaata tgactacaag 180
ctcctggccg acgattcgaa ggagaaggtg gcgagccacg tacgaggatg tgcctgccac 240
ctcaggccat gcatccggtt ttgttgcccc cagtaccaaa agatgcaaaa gagcaagtgc 300
tatggcgaca tgtcggagga cgagctgaac aagcacgatc ccttcgtgaa cgtgacgctc 360
agcgacgggt cggtggtcag gagacacttc aaggaggatc tgatcgtgca gtcggatctg 420
gccaagcccg gatgtccccg gatgtacttt ttgaatcacg aactgccggg caatgaattc 480
actctgtttg agaacggctc tcttctgcgt cactgggaca aagtggaact gagcaagcgg 540
gagtactgcg tccagcatct ttcatttaag gatgatagta tccgaattgc accccacttc 600
tgtccacttt cgtctgagca ctccagaacg tggaagaccg ttgcgatagt gatatccttg 660
atatgcataa tcctaacgat cagcgtgtac ctctatgtcg agaagcttcg caacctgcac 720
ggaaagtgtt tcatctgcta cttggcctct ttgttcttgg gctatttctt cctggtcctc 780
aacgtgtgga aatactcctc cggcttctgc gttacagcag ggttcctggg ctacttctcc 840
gtcatggccg cgttcttttg gctttctgtg attggaatcc acttgcggat taagtttagt 900
ctagcctcta actgtttgca ccgattgctg ccagaaaatc cgttccgtgc atacaactta 960
tacgcctggg gtataccgct aatcatgact gcaatcacct atacagctga ccaggtggtt 1020
aagaacgaga agttgagacc ccgcgtgggc gtcggcaaaa attgttggat ctacacaagc 1080
ttgggctag 1089

<210> 57
<211> 362
<212> PRT
<213> Drosophila

<400> 57
Met Arg Ile Leu Leu Ile Ala Val Leu Phe Leu Leu Met Pro Lys Ser
1 5 10 15
Asn Ala Glu Ile Pro Gly Cys Asp Phe Phe Asp Thr Val Asp Ile Ser
20 25 30
Lys Ala Pro Arg Phe Ser Asn Gly Ser Tyr Leu Tyr Glu Gly Leu Leu
35 40 45
Ile Pro Ala His Leu Thr Ala Glu Tyr Asp Tyr Lys Leu Leu Ala Asp
50 55 60
Asp Ser Lys Glu Lys Val Ala Ser His Val Arg Gly Cys Ala Cys His
65 70 75 80
Leu Arg Pro Cys Ile Arg Phe Cys Cys Pro Gln Tyr Gln Lys Met Gln
85 90 95
Lys Ser Lys Cys Tyr Gly Asp Met Ser Glu Asp Glu Leu Asn Lys His
100 105 110
Asp Pro Phe Val Asn Val Thr Leu Ser Asp Gly Ser Val Val Arg Arg
115 120 125
His Phe Lys Glu Asp Leu Ile Val Gln Ser Asp Leu Ala Lys Pro Gly
130 135 140
Cys Pro Arg Met Tyr Phe Leu Asn His Glu Leu Pro Gly Asn Glu Phe
145 150 155 160
Thr Leu Phe Glu Asn Gly Ser Leu Leu Arg His Trp Asp Lys Val Glu
165 170 175
Leu Ser Lys Arg Glu Tyr Cys Val Gln His Leu Ser Phe Lys Asp Asp

```
            180                 185                 190
Ser Ile Arg Ile Ala Pro His Phe Cys Pro Leu Ser Ser Glu His Ser
        195                 200                 205
Arg Thr Trp Lys Thr Val Ala Ile Val Ile Ser Leu Ile Cys Ile Ile
    210                 215                 220
Leu Thr Ile Ser Val Tyr Leu Tyr Val Glu Lys Leu Arg Asn Leu His
225                 230                 235                 240
Gly Lys Cys Phe Ile Cys Tyr Leu Ala Ser Leu Phe Leu Gly Tyr Phe
                245                 250                 255
Phe Leu Val Leu Asn Val Trp Lys Tyr Ser Ser Gly Phe Cys Val Thr
            260                 265                 270
Ala Gly Phe Leu Gly Tyr Phe Ser Val Met Ala Ala Phe Phe Trp Leu
        275                 280                 285
Ser Val Ile Gly Ile His Leu Arg Ile Lys Phe Ser Leu Ala Ser Asn
    290                 295                 300
Cys Leu His Arg Leu Leu Pro Glu Asn Pro Phe Arg Ala Tyr Asn Leu
305                 310                 315                 320
Tyr Ala Trp Gly Ile Pro Leu Ile Met Thr Ala Ile Thr Tyr Thr Ala
                325                 330                 335
Asp Gln Val Val Lys Asn Glu Lys Leu Arg Pro Arg Val Gly Val Gly
            340                 345                 350
Lys Asn Cys Trp Ile Tyr Thr Ser Leu Gly
        355                 360
```

<210> 58
<211> 7503
<212> DNA
<213> Drosophila

<400> 58

```
gattgccgtg gtcgtgcgcc gcgggagaaa tgcccttttc gtgtacgcaa ttactaggca 60
aaaaggtttg gacggccctt gctttattag gccataattt acgctgtcga gcggccatta 120
gagcggcatc cgctggccaa gttggtctat aaaagcccgg cggctgcaag tggcgttttc 180
cattccacgt gaaaggacac tcgcagtcgg gccgacgaga tgctgcgcca cctgctccgc 240
cacgagaaca acaaggtctt cgtcctgatc ctgctctact gcgtcctggt cagcattctg 300
aaactctgca cggcacagcc ggatagctct gtggccgcca cggataatgg tatttgataa 360
ggcatttgat aatatcatat tattcgctct aatggattcc atttaccatc cagatattac 420
gcatcttggc gacgattgtc aggtgacgcc cgtcatccat gtgctccagt atcctggatg 480
tgtgcccaag ccgattccct cgttcgcctg cgtgggtcgc tgtgccagtt atatccaggt 540
gggtgttgat atacttgaaa tataaattat aatttgatat aatctgatga ctttaggttt 600
cgggcagtaa gatctggcaa atggagcgtt cctgcatgtg ctgccaggag tctggtgagc 660
gggaggcagc cgtctcgcta ttctgtccca aagtgaagcc cggcgagcgt aaattcaaga 720
aggtcctgac caaggcgcca ttggagtgca tgtgtcggcc atgcacttcc attgaggagt 780
ctggcatcat accacaggaa attgccggct attcggacga gggtccactc aacaatcact 840
tccggcgcat tgctctgcaa tagattcccc catcagttta gcactcatac ccatgcccat 900
ttgctcatta aaataaatta gagttgccgt ttctgcttgt acacagtgaa aatgtgcct 960
ctgaaagctc atgccagtga aagggaaaca tcagcagata tatgattatt ttaaattgta 1020
atttatttcc caactatctc atgtcattgc tactagttca ctaaatggtg tgtgtcctgc 1080
ttcacagggt atgtatgtag atgcgtaact ccaaagaaaa caaagataat ctttacactg 1140
caactggatc taacaatagg agtttaacta cggcataatg catatgactc gcgccgccta 1200
caagtactcg acgatatcgc atttgatgga tatattcgga cagctctggc ccacggcact 1260
gcggccccga tagttgggac tcttctcggt gaggagtttc cggaagatgg gcaaaattgc 1320
cgtctcgccc tcgagcagtt ccgtctcgct gctggcgccg tagatgctgc tatcctgctt 1380
gccgcgctgc ctgctccgct gcttatcggg cgccacattg ggcgaactgt ggtggtgggc 1440
ggcgtgctgg atgttagagc tccggctgac cttacgcggc tggacgcacg gctgactgtt 1500
gggcggcgga tctcgttccc gttcgctgtc cgtctcgctc tggcaaacgg tgcccttttg 1560
cggttttggc cgatgtttgt tgcggttctt ctcggagacg gaccgctgga tgggcgaggt 1620
ggtgttcggg tcataggtca tcccgtggcc gccatgatgt tccgcacagc gctgcggact 1680
ggcgtagttc cgcggcgact ggcgctgctt ggcgtgcgac agcgtgtgcg tgtcccagag 1740
```

gtcactctgt gaaaccacga aattggactt cagagcggag cccatgctgg tgcggcgatg 1800
gtcggaggca gtgggagcta ttttcgatta gaaaatcgat agtattatag tattataaga 1860
gacgcgcaga aaacttactc ttaatcacag ccggagtact gggctttgag tgtggacaca 1920
tgttgctcaa ctccgattgc gatgtggaga tgttggtgga ggagttccgg tgcgagctgc 1980
tgtactttgg ccgggcggtg ggcgtctggc tgccactgtc gtagtgcctg gaggcggtcg 2040
attgacgcct ctcccgcatg ctggagcagt ctcctccgct ccgcttgccc gccgtactgg 2100
tgatgctgcc tatattatca tcggtctgag cggacaccga gtttctcttc tccagatggc 2160
gatgatgatg ctgctgctgc tgctgttgct gttgcatctg ctggtgttgc tgctgctgca 2220
ggtgctgctg ctgcatctgc tgatgttgct gtaagtgctg ctgcagttgc tgctggatgg 2280
gctgcatcat agttggagca ggggcatgcg agtggtgctc tacaatcgat tgcttttct 2340
ttttgctgtt ggagaatatc catagattag agtgtattta agaccaaaat ttaagaactt 2400
acggaggaag cgatggcgcc atcactgtag actgcaccga cacaaaggag ttatccatct 2460
ccacatgagt cgaggtcaca ctatccacgg acgtcatctc ggtaacttct atgacaaacg 2520
ggttgcaatg catgaacata tctacgaaat ccccctcctt agccactcac ctgtggtgct 2580
gggcatctct cgacgcaccg gtggagctga gagcctaaca atgtcatcgt tcagtatggg 2640
ctccagttcg ctaccgacgt tggagccacc tgtgcaggtc accccatcga tccatttgcg 2700
tgcctcgggt cccagcttgc ggattaaggc ctgcagctcg ttctccttct ccatcagcgc 2760
ctttcggaat cggcagtttg tgttctttac atctcgcaat cgttgctcca gttcgcacac 2820
cgaggaatcc cggcgtccgt ttttggacat gggtctcaac gtggccctaa cgcgtttgtc 2880
caccacgccc tggggattcc gcttcagctc caccaactgc aaaacgattc gcatgaatat 2940
taaccaaaga ttatactacg cggtcaacat ctaggaagga tcaactctac ggacagtaat 3000
ataggatagt taaagatatt ccaacagcaa ggaagtggga tttgcctaca accttgctgc 3060
tgctcttagt gggttctata tagttcaagt ttatagttag tctgtagtgt tatatagcat 3120
tttaagatct gttttcgagt acatgttgtt tctgtatcaa ggtgtatcaa cgctagacga 3180
aagatcagct cagctggcgg taaagtacga tcagttaaga agctctaacc acaactgatc 3240
ggcctattac cgccgacaaa caacattcat attacacgat tattatatgt aacatataac 3300
agctacaaaa actacaaaac ttccttacta ccagtcttac tttcggtacg aacaccaaac 3360
acaaagtggc tgtcgtacaa aaaatgataa aaaacgagag taagacaaaa actaaatcct 3420
ttcgatccga tagcaccagg gatatagccg ctccggccag acaagtgatg aacacgttat 3480
aaacggagaa accaatatgc ttggaatcgt tcagagccgg tatagaaaca tgtcgagttt 3540
cccaggccaa aaaggcgcca aaaacctatt gaaaaaaatt attgctttat tataatgata 3600
tcttacaaat taatgtttat tatatactca ccaacaacag tcccttgtag gcataaataa 3660
tgctaacgaa tatggtcatg tgctcagact ggcagtactc gttttcgggg atcaccaaga 3720
catcatcaat attctcgtga tgctgtagtt aataataatc attttagatt tttcacttca 3780
aatatgttct atcaatgcac tcaccaaggg ttccaactgt ttagtttcgc ggtaaaaggg 3840
atcggcaatc tgccaggtgg ttataatggc tatatcaatg gccaaaagca cgcccacaac 3900
cataaacaat tgatagtcct tgatcacctt cttattgagc ttcagatcgg tgaatatcga 3960
atgcacccgc cacgtcttcg agaacatggc tccaaaactg agactgaatc cagccatcaa 4020
gatccaggct cgagctgtgc agatataggg aaaagctgcc acactactta atgtggtatc 4080
gagacccagg aaaatgatgc tcaaataggt catcatacag cccacaatga tcagattgtt 4140
caaatgggga ctggacatct tgatgtatct ataaaaacat caatatttta ttaaacgaaa 4200
ttgtttaata tagcattgtt aagctcacct ttgattgcga tacttaatgt taaaggccag 4260
aaaaactgtg gcaataatca ctccaatgac cgaagcacta gccgatacaa tatatatggt 4320
tggattgacc tgactgtgct cgatgtagat caaagtgcga tccttgggag gagttttccc 4380
cacccatttg actggttttc ccaagcttaa atccaagtgt gacttctgtg agtggtattc 4440
cccgatcttt tccatttgtc ccagctgaaa ctgattgatc aggatgttgg ccttgcgctc 4500
gttgttgtag aaacgcacgg gtcccttgag aatgtaaaat atgttttaac tgcaaaaact 4560
ggattgttta tctatctacc gtcacaccct cgaaggatgt attacgtaga gcctcaagga 4620
agacactctc ccagtccttc acgcgataat caaaatgtgt tagcagatcc tctcgctttt 4680
cggccacata ctgaatggcc agggcagctg cccagatacc atcgtaggta tagccatgaa 4740
agcgggaata ttcagtgcct cgcagtctgt cgtactccac aagatactca tcagcagtct 4800
agtaaaatat gaaatttatt acaagaattt taatagactt tactggtact aacaatgcca 4860
gccactgtga tgtcaccact ggtggacaag ggcaaaagat ccactagaat ggcaccttcc 4920
aaggctgtag cgatctcctc cacactgcac tcgctgtcct gcgtgacatt ccaccaatcc 4980
gtggaatagg tagccatgat cagccattga taggctctgc catacatatc caatctggaa 5040
gggagtttat tgcaataatt ttaaattgag acttaagtta gtatatccaa ctcacttata 5100
agcctcacag aatgccttgc gtgcaaagtg ctcgttaaag ttgcccagaa tgatcctcac 5160
gtccttctcg cgcagtttct tcaatgattc agccacatcg ttgacgaagc tctgcgtttc 5220
caccacctcg acctccatgg catccaggtc agccaccatg tgattgtggg gcagcgaata 5280
gcgtggctca ttctggtaga cagtgcccac tctggtccaa ttgaactcct tcagcaaggc 5340

cagtcgcggc gcattaaagg cattctccga gggtaccacg cgaaagaaat tcggaaacgc 5400
atccttggtg aacatgggat gggtgtccgc gtagctgagc tgggtgaggt gccagtgctt 5460
gctggccttg gctatgggat cggtcacatg ggtgcacgca gcgccgaaga gcatcacttt 5520
attgggaccc gaatgcatca tatcgaagaa ggactttacg cccacagcag cattgcactg 5580
gaaaggaaac caaaattaca ttccagaata tatggaaata tgcataggta aaggtgtgat 5640
attacaatat gaaccaatga acattaaaaa aaatacctaa ttgaacttta aaaaccacaa 5700
gtccatcatc agagattcac ttaatgaaat cttcaagcta aaccctaact cacctgagtg 5760
tcgttccacc acatgtgcag cctgtagttg gccagtatct ttccatgctc attaacgtga 5820
cccaaggcga gctttacact gggcataacg ccccgaccgg tgtaggagtt ttccacgcca 5880
tccccgtacg ggaagaatcc cgctatgtag acgtccgatc tcttggctgt cctgccacag 5940
gcggtgctcc aaaggagcag gaacagcagg ctggcgaatg gaaaccaact tggccggaac 6000
atgttgaccg cgtgggctgt aaaacgattc taaatgtgaa tagcgcatat cgaggacgtg 6060
aacttggagc caagttgtgg gtggtgaatg gcagtggttg gtttcgggtg agccaaggct 6120
gccgaggcaa ggctgcgccc agtgggtgtg tattattttc tacgttctac ctgtgtttgc 6180
tggcactcgc atgcaatttt acaggacagg tggagcaagg ctaaggccag tccaggccag 6240
caaaattgct tcaaattgct gcagcaatgc tggcacgttt agtaattata aagcaattta 6300
atgatctaaa ggattcgtgc attatttagt atgctcgtta acgcgaagga gtatcttcca 6360
ggtgttgcat cctcttaaat cttatgtgta catgttttta ggaacctacc agcttgctta 6420
gtcccagaac tggcttttat ttcactaccg atgccatttg aaacacttat tttacatcaa 6480
atctcagcca acgaagaact tttgcagtgc gagggtccga aaaattatgt atccccctga 6540
cttgacattt gtctgacagg aggaattcga ggcaggaagt gcgcatgtcc ttaaaccgaa 6600
agctgtcact gctccgcctt ggcctaggct ttccaacccc cgtaattacc ataaacccac 6660
tgcgcttgta aacatcctgc aacaggtggt cccatcagct ggtactccgg caattgttgt 6720
gcccgttgcc atggactact ggggtcgaaa aacgaagtag ataactgctc ggaaaattac 6780
caaattgaaa attgtgagta accataatgg aggactatgc gcccagccgc tataaaatgc 6840
ttgagaaaat cggcgaaggc gtccatggct gcgtgtttaa ggccatcgat ctgcagcgca 6900
acaaggaggt ggccatcaag aaggtggccc ttaagaacaa gttcggaaac atagccctaa 6960
atactttgag ggaaatcaag accctgcagc tttgcaagtc tgaatatgta ggtataccat 7020
atatctattt tccatttgtt aattatgtaa ctaattgtag attctggaca tcattgatat 7080
atatcctgac ctaactggtc tctccttggt gctggaatac cagccggaca ctctgtacaa 7140
ccgcctcaag agcgaggtta atcccctgag ccgccagcag gtgcgtaagt ttgcacacca 7200
aatgttcaag ggattgcat atctgcatga agcgggtctc atgcaccggg atattaaacc 7260
cgccaatctg cttatcagcg atacagatat gcttaagatt gccgatttcg gactggcccg 7320
tctgtacttc cctgaagacg agtcgcgcct ttattcgcct caggtctcaa cccggtggta 7380
tcgagcgcca gagattctat ttggcagcca gaagtatgga accggagtgg atatgtgggc 7440
ggcaggttag ccatgactct ttgatgggtt tacatatatc cttatggctg acattattgc 7500
tat 7503

<210> 59
<211> 3971
<212> DNA
<213> Drosophila

<400> 59
aaaagttctt cgttggctga gatttgatgt aaaataagtg tttcaaatgg catcggtagt 60
gaaataaaag ccagttctgg gactaagcaa gctgcccacg cggtcaacat gttccggcca 120
agttggtttc cattcgccag cctgctgttc ctgctccttt ggagcaccgc ctgtggcagg 180
acagccaaga gatcggacgt ctacatagcg ggattcttcc cgtacgggga tggcgtggaa 240
aactcctaca ccggtcgggg cgttatgccc agtgtaaagc tcgccttggg tcacgttaat 300
gagcatggaa agatactggc caactacagg ctgcacatgt ggtggaacga cactcagtgc 360
aatgctgctg tgggcgtaaa gtccttcttc gatatgatgc attcgggtcc caataaagtg 420
atgctcttcg gcgctgcgtg cacccatgtg accgatccca tagccaaggc cagcaagcac 480
tggcacctca cccagctcag ctacgcggac acccatccca tgttcaccaa ggatgcgttt 540
ccgaatttct ttcgcgtggt accctcggag aatgccttta atgcgccgcg actggccttg 600
ctgaaggagt tcaattggac cagagtgggc actgtctacc agaatgagcc acgctattcg 660
ctgccccaca atcacatggt ggctgacctg gatgccatgg aggtcgaggt ggtggaaacg 720
cagagcttcg tcaacgatgt ggctgaatca ttgaagaaac tgcgcgagaa ggacgtgagg 780
atcattctgg gcaactttaa cgagcacttt gcacgcaagg cattctgtga ggcttataaa 840

ttggatatgt atggcagagc ctatcaatgg ctgatcatgg ctacctattc cacggattgg 900
tggaatgtca cgcaggacag cgagtgcagt gtggaggaga tcgctacagc cttggaaggt 960
gccattctag tggatctttt gcccttgtcc accagtggtg acatcacagt ggctggcatt 1020
actgctgatg agtatcttgt ggagtacgac agactgcgag gcactgaata ttcccgcttt 1080
catggctata cctacgatgg tatctgggca gctgccctgg ccattcagta tgtggccgaa 1140
aagcgagagg atctgctaac acattttgat tatcgcgtga aggactggga gagtgtcttc 1200
cttgaggctc tacgtaatac atccttcgag ggtgtgacgg gacccgtgcg tttctacaac 1260
aacgagcgca aggccaacat cctgatcaat cagtttcagc tgggacaaat ggaaaagatc 1320
ggggaatacc actcacagaa gtcacacttg gatttaagct tgggaaaacc agtcaaatgg 1380
gtggggaaaa ctcctcccaa ggatcgcact ttgatctaca tcgagcacag tcaggtcaat 1440
ccaaccatat atattgtatc ggctagtgct tcggtcattg gagtgattat tgccacagtt 1500
tttctggcct ttaacattaa gtatcgcaat caaagataca tcaagatgtc cagtccccat 1560
ttgaacaatc tgatcattgt gggctgtatg atgacctatt tgagcatcat tttcctgggt 1620
ctcgatacca cattaagtag tgtggcagct ttccctata tctgcacagc tcgagcctgg 1680
atcttgatgg ctggattcag tctcagtttt ggagccatgt tctcgaagac gtggcgggtg 1740
cattcgatat tcaccgatct gaagctcaat aagaaggtga tcaaggacta tcaattgttt 1800
atggttgtgg gcgtgctttt ggccattgat atagccatta taaccacctg gcagattgcc 1860
gatccctttt accgcgaaac taaacagttg gaacccttgc atcacgagaa tattgatgat 1920
gtcttggtga tccccgaaaa cgagtactgc cagtctgagc acatgaccat attcgttagc 1980
attatttatg cctacaaggg actgttgttg gtttttggcg cctttttggc ctgggaaact 2040
cgacatgttt ctataccggc tctgaacgat tccaagcata ttggtttctc cgtttataac 2100
gtgttcatca cttgtctggc cggagcggct atatccctgg tgctatcgga tcgaaaggat 2160
ttagtttttg tcttactctc gtttttatc atttttgta cgacagccac tttgtgtttg 2220
gtgttcgtac cgaaattggt ggagctgaag cggaatcccc agggcgtggt ggacaaacgc 2280
gttagggcca cgttgagacc catgtccaaa aacggacgcc gggattcctc ggtgtgcgaa 2340
ctggagcaac gattgcgaga tgtaaagaac acaaactgcc gattccgaaa ggcgctgatg 2400
gagaaggaga acgagctgca ggccttaatc cgcaagctgg gacccgaggc acgcaaatgg 2460
atcgatgggg tgacctgcac aggtggctcc aacgtcggta gcgaactgga gcccatactg 2520
aacgatgaca ttgttaggct ctcagctcca ccggtgcgtc gagagatgcc cagcaccaca 2580
gaagttaccg agatgacgtc cgtggatagt gtgacctcga ctcatgtgga gatggataac 2640
tcctttgtgt cggtgcagtc tacagtgatg gcgccatcgc ttcctcccaa aaagaaaaag 2700
caatcgattg tagagcacca ctcgcatgcc cctgctccaa ctatgatgca gcccatccag 2760
cagcaactgc agcagcactt acagcaacat cagcagatgc agcagcagca cctgcagcag 2820
cagcaacacc agcagatgca acagcaacag cagcagcagc agcatcatca tcgccatctg 2880
gagaagagaa actcggtgtc cgctcagacc gatgataata taggcagcat caccagtacg 2940
gcgggcaagc ggagcggagg agactgctcc agcatgcggg agaggcgtca atcgaccgcc 3000
tccaggcact acgacagtgg cagccagacg cccaccgccc ggccaaagta cagcagctcg 3060
caccggaact cctccaccaa catctccaca tcgcaatcgg agttgagcaa catgtgtcca 3120
cactcaaagc ccagtactcc ggctgtgatt aagactccca ctgcctccga ccatcgccgc 3180
accagcatgg gctccgctct gaagtccaat ttcgtggttt cacagagtga cctctgggac 3240
acgcacacgc tgtcgcacgc caagcagcgc cagtcgccgc ggaactacgc cagtccgcag 3300
cgctgtgcgg aacatcatgg cggccacggg atgacctatg acccgaacac cacctcgccc 3360
atccagcggt ccgtctccga gaagaaccgc aacaaacatc ggccaaaacc gcaaaagggc 3420
accgtttgcc agagcgagac ggacagcgaa cgggaacgag atccgccgcc caacagtcag 3480
ccgtgcgtcc agccgcgtaa ggtcagccgg agctctaaca tccagcacgc cgcccaccac 3540
cacagttcgc ccaatgtggc gcccgataag cagcggagca ggcagcgcgg caagcaggat 3600
agcagcatct acggcgccag cagcgagacg gaactgctcg agggcgagac ggcaattttg 3660
cccatcttcc ggaaactcct caccgagaag agtcccaact atcggggccg cagtgccgtg 3720
ggccagagct gtccgaatat atccatcaaa tgcgatatcg tcgagtactt gtaggcggcg 3780
cgagtcatat gcattatgcc gtagttaaac tcctattgtt agatccagtt gcagtgtaaa 3840
gattatcttt gtttctttg gagttacgca tctacataca taccctgtga agcaggacac 3900
acaccattta gtgaactagt agcaatgaca tgagatagtt gggaaataaa ttacaattta 3960
aaataatcat a 3971

<210> 60
<211> 1221
<212> PRT
<213> Drosophila

<400> 60

```
Met Phe Arg Pro Ser Trp Phe Pro Phe Ala Ser Leu Leu Phe Leu Leu
 1               5                  10                  15
Leu Trp Ser Thr Ala Cys Gly Arg Thr Ala Lys Arg Ser Asp Val Tyr
            20                  25                  30
Ile Ala Gly Phe Phe Pro Tyr Gly Asp Gly Val Glu Asn Ser Tyr Thr
        35                  40                  45
Gly Arg Gly Val Met Pro Ser Val Lys Leu Ala Leu Gly His Val Asn
    50                  55                  60
Glu His Gly Lys Ile Leu Ala Asn Tyr Arg Leu His Met Trp Trp Asn
65                  70                  75                  80
Asp Thr Gln Cys Asn Ala Ala Val Gly Val Lys Ser Phe Phe Asp Met
                85                  90                  95
Met His Ser Gly Pro Asn Lys Val Met Leu Phe Gly Ala Ala Cys Thr
            100                 105                 110
His Val Thr Asp Pro Ile Ala Lys Ala Ser Lys His Trp His Leu Thr
        115                 120                 125
Gln Leu Ser Tyr Ala Asp Thr His Pro Met Phe Thr Lys Asp Ala Phe
    130                 135                 140
Pro Asn Phe Phe Arg Val Val Pro Ser Glu Asn Ala Phe Asn Ala Pro
145                 150                 155                 160
Arg Leu Ala Leu Leu Lys Glu Phe Asn Trp Thr Arg Val Gly Thr Val
                165                 170                 175
Tyr Gln Asn Glu Pro Arg Tyr Ser Leu Pro His Asn His Met Val Ala
            180                 185                 190
Asp Leu Asp Ala Met Glu Val Glu Val Val Glu Thr Gln Ser Phe Val
        195                 200                 205
Asn Asp Val Ala Glu Ser Leu Lys Lys Leu Arg Glu Lys Asp Val Arg
    210                 215                 220
Ile Ile Leu Gly Asn Phe Asn Glu His Phe Ala Arg Lys Ala Phe Cys
225                 230                 235                 240
Glu Ala Tyr Lys Leu Asp Met Tyr Gly Arg Ala Tyr Gln Trp Leu Ile
                245                 250                 255
Met Ala Thr Tyr Ser Thr Asp Trp Trp Asn Val Thr Gln Asp Ser Glu
            260                 265                 270
Cys Ser Val Glu Glu Ile Ala Thr Ala Leu Glu Gly Ala Ile Leu Val
        275                 280                 285
Asp Leu Leu Pro Leu Ser Thr Ser Gly Asp Ile Thr Val Ala Gly Ile
    290                 295                 300
Thr Ala Asp Glu Tyr Leu Val Glu Tyr Asp Arg Leu Arg Gly Thr Glu
305                 310                 315                 320
Tyr Ser Arg Phe His Gly Tyr Thr Tyr Asp Gly Ile Trp Ala Ala Ala
                325                 330                 335
Leu Ala Ile Gln Tyr Val Ala Glu Lys Arg Glu Asp Leu Leu Thr His
            340                 345                 350
Phe Asp Tyr Arg Val Lys Asp Trp Glu Ser Val Phe Leu Glu Ala Leu
        355                 360                 365
Arg Asn Thr Ser Phe Glu Gly Val Thr Gly Pro Val Arg Phe Tyr Asn
    370                 375                 380
Asn Glu Arg Lys Ala Asn Ile Leu Ile Asn Gln Phe Gln Leu Gly Gln
385                 390                 395                 400
Met Glu Lys Ile Gly Glu Tyr His Ser Gln Lys Ser His Leu Asp Leu
                405                 410                 415
Ser Leu Gly Lys Pro Val Lys Trp Val Gly Lys Thr Pro Pro Lys Asp
            420                 425                 430
Arg Thr Leu Ile Tyr Ile Glu His Ser Gln Val Asn Pro Thr Ile Tyr
        435                 440                 445
Ile Val Ser Ala Ser Ala Ser Val Ile Gly Val Ile Ile Ala Thr Val
    450                 455                 460
Phe Leu Ala Phe Asn Ile Lys Tyr Arg Asn Gln Arg Tyr Ile Lys Met
465                 470                 475                 480
```

Ser Ser Pro His Leu Asn Asn Leu Ile Ile Val Gly Cys Met Met Thr
485 490 495
Tyr Leu Ser Ile Ile Phe Leu Gly Leu Asp Thr Thr Leu Ser Ser Val
500 505 510
Ala Ala Phe Pro Tyr Ile Cys Thr Ala Arg Ala Trp Ile Leu Met Ala
515 520 525
Gly Phe Ser Leu Ser Phe Gly Ala Met Phe Ser Lys Thr Trp Arg Val
530 535 540
His Ser Ile Phe Thr Asp Leu Lys Leu Asn Lys Lys Val Ile Lys Asp
545 550 555 560
Tyr Gln Leu Phe Met Val Val Gly Val Leu Leu Ala Ile Asp Ile Ala
565 570 575
Ile Ile Thr Thr Trp Gln Ile Ala Asp Pro Phe Tyr Arg Glu Thr Lys
580 585 590
Gln Leu Glu Pro Leu His His Glu Asn Ile Asp Asp Val Leu Val Ile
595 600 605
Pro Glu Asn Glu Tyr Cys Gln Ser Glu His Met Thr Ile Phe Val Ser
610 615 620
Ile Ile Tyr Ala Tyr Lys Gly Leu Leu Leu Val Phe Gly Ala Phe Leu
625 630 635 640
Ala Trp Glu Thr Arg His Val Ser Ile Pro Ala Leu Asn Asp Ser Lys
645 650 655
His Ile Gly Phe Ser Val Tyr Asn Val Phe Ile Thr Cys Leu Ala Gly
660 665 670
Ala Ala Ile Ser Leu Val Leu Ser Asp Arg Lys Asp Leu Val Phe Val
675 680 685
Leu Leu Ser Phe Phe Ile Ile Phe Cys Thr Thr Ala Thr Leu Cys Leu
690 695 700
Val Phe Val Pro Lys Leu Val Glu Leu Lys Arg Asn Pro Gln Gly Val
705 710 715 720
Val Asp Lys Arg Val Arg Ala Thr Leu Arg Pro Met Ser Lys Asn Gly
725 730 735
Arg Arg Asp Ser Ser Val Cys Glu Leu Glu Gln Arg Leu Arg Asp Val
740 745 750
Lys Asn Thr Asn Cys Arg Phe Arg Lys Ala Leu Met Glu Lys Glu Asn
755 760 765
Glu Leu Gln Ala Leu Ile Arg Lys Leu Gly Pro Glu Ala Arg Lys Trp
770 775 780
Ile Asp Gly Val Thr Cys Thr Gly Gly Ser Asn Val Gly Ser Glu Leu
785 790 795 800
Glu Pro Ile Leu Asn Asp Asp Ile Val Arg Leu Ser Ala Pro Pro Val
805 810 815
Arg Arg Glu Met Pro Ser Thr Thr Glu Val Thr Glu Met Thr Ser Val
820 825 830
Asp Ser Val Thr Ser Thr His Val Glu Met Asp Asn Ser Phe Val Ser
835 840 845
Val Gln Ser Thr Val Met Ala Pro Ser Leu Pro Pro Lys Lys Lys Lys
850 855 860
Gln Ser Ile Val Glu His His Ser His Ala Pro Ala Pro Thr Met Met
865 870 875 880
Gln Pro Ile Gln Gln Gln Leu Gln Gln His Leu Gln Gln His Gln Gln
885 890 895
Met Gln Gln Gln His Leu Gln Gln Gln Gln His Gln Gln Met Gln Gln
900 905 910
Gln Gln Gln Gln Gln Gln His His His Arg His Leu Glu Lys Arg Asn
915 920 925
Ser Val Ser Ala Gln Thr Asp Asp Asn Ile Gly Ser Ile Thr Ser Thr
930 935 940
Ala Gly Lys Arg Ser Gly Gly Asp Cys Ser Ser Met Arg Glu Arg Arg
945 950 955 960

Gln Ser Thr Ala Ser Arg His Tyr Asp Ser Gly Ser Gln Thr Pro Thr
965 970 975
Ala Arg Pro Lys Tyr Ser Ser Ser His Arg Asn Ser Ser Thr Asn Ile
980 985 990
Ser Thr Ser Gln Ser Glu Leu Ser Asn Met Cys Pro His Ser Lys Pro
995 1000 1005
Ser Thr Pro Ala Val Ile Lys Thr Pro Thr Ala Ser Asp His Arg Arg
1010 1015 1020
Thr Ser Met Gly Ser Ala Leu Lys Ser Asn Phe Val Val Ser Gln Ser
1025 1030 1035 1040
Asp Leu Trp Asp Thr His Thr Leu Ser His Ala Lys Gln Arg Gln Ser
1045 1050 1055
Pro Arg Asn Tyr Ala Ser Pro Gln Arg Cys Ala Glu His His Gly Gly
1060 1065 1070
His Gly Met Thr Tyr Asp Pro Asn Thr Thr Ser Pro Ile Gln Arg Ser
1075 1080 1085
Val Ser Glu Lys Asn Arg Asn Lys His Arg Pro Lys Pro Gln Lys Gly
1090 1095 1100
Thr Val Cys Gln Ser Glu Thr Asp Ser Glu Arg Glu Arg Asp Pro Pro
1105 1110 1115 1120
Pro Asn Ser Gln Pro Cys Val Gln Pro Arg Lys Val Ser Arg Ser Ser
1125 1130 1135
Asn Ile Gln His Ala Ala His His His Ser Ser Pro Asn Val Ala Pro
1140 1145 1150
Asp Lys Gln Arg Ser Arg Gln Arg Gly Lys Gln Asp Ser Ser Ile Tyr
1155 1160 1165
Gly Ala Ser Ser Glu Thr Glu Leu Leu Glu Gly Glu Thr Ala Ile Leu
1170 1175 1180
Pro Ile Phe Arg Lys Leu Leu Thr Glu Lys Ser Pro Asn Tyr Arg Gly
1185 1190 1195 1200
Arg Ser Ala Val Gly Gln Ser Cys Pro Asn Ile Ser Ile Lys Cys Asp
1205 1210 1215
Ile Val Glu Tyr Leu
1220

<210> 61
<211> 18996
<212> DNA
<213> Drosophila

<400> 61
gttccttaaa cgtgatcaaa aatataacta gagctactac taagcagcga gcacagcgtg 60
cgatatatca acaagtccag ctggcaacat gttgccgcgc ctgtggtgag ttgtgcctct 120
atgtcctttg gcagcatgtt gcggtcgctt ggcgccggct ggtggtcaca aaaagcgccg 180
acgcttgccg tcagtgcttc gcgaagattg cgaggcgcga cacacatcgc ggaacccgaa 240
ctccatacgg ctgtgccgac actcaatccc gacctaagcc caattttttg ttgcgtgaag 300
ttgacgaagt gtgttgtact gctgcaagtg caagtccccg attgctagtg ctattccggg 360
atcacaaacg agcatccttc ttcggcctca cgattgacgc cttctaccac tacctgcgcc 420
aggcgctgcc gctggccaag gaggctgcga tccacctgaa tgccagcaac gagatctccg 480
ctgtcggtga cggggtcacc atcacaggga cacccggtga tcttctcaac tacagcgggc 540
tggaattgga tctcggactg gatttggatc ttaatctgga catggaccta gccactactc 600
ccagcagcag caccttggcg ccagctgtaa ccgtcaggac accggggaac aggagcgtgg 660
tgagggtcag tgccgatgtg cccatctggg tggtgccgtg ctattcggcc atccttctgt 720
gcgccgtggt cgggaacctg ctggtggtgc taacgctcgt ccagaatcgt cgcatgcgca 780
ccatcacgaa cgtttttctg ctcaacctgg ccatctcgga catcctgctc ggcgtattct 840
gcatgccggt gaccctggtc ggcactctgc taaggcattt catattcggc gagctgctct 900
gcaagctcat ccagtttgct cagggtgagt ggtggaagcc agtaggccag ttcatttatg 960
tcttgtccgt gtccaaatta aaaggttatc ccaaacttcc aactgtgaat tccgcattta 1020
ttagtcccta tccccgtgac cgagccccat aatcgcttgt caaatggcca cgaaccagca 1080

cgacctcgag cggattcgcc agcagatcgt gttggccaac attcaggagc tgataaagaa 1140
gatgacacgt cgctgcttcg acgtatgtat cgctatgccg gaaatggagt tgcgctccac 1200
ggagcgcgac tgcctggcca actgtatgga tcgattcatg gactcggttc aggtggtgtc 1260
gagccagtac ttccgtcgcc ggcgtcgcca tcagcaaatt cgtttgtccc gttcgaccgc 1320
ctcctccgca tcaccaccag catccgcatc cgcatccatg cccaaatccg cagcagcaaa 1380
tgaatctgaa tccgcatcta gggcctccaa tgacgaaaag gtcaaataat ggttaagtct 1440
agacagatag ccacgtgtgg agtttattcc aagtcgaaaa cgaaggctgt tccgtcgtag 1500
ccccagatcc ccagtgtact tgggtgatgg ccacttgtcc acattggtca tcatcccaaa 1560
gctttcgagg atgatttcat acttcgagac ctttgggaag ggcaagaagg ggtcgtctag 1620
ctggaaatga aagtggattt tctgcaggcg tggccatcga tgatttcttg ctattggttc 1680
ttaacaaaat cctagtagag tcttaagtaa atgacgattg gcagaaaata aagatgtcgt 1740
tcagtgtaca ttcactatgc tttgagattc tctttaaagt caccctttttg aaatcactta 1800
gctaagctta atattgtata atatacaaag tagtgtagtg gggagaagat aagttttaaa 1860
tgaccatact gtatatttta gattgtgcgc ttgtgcacat cacttttcca cttaggaaaa 1920
ctgagtctgc catcgatttt acgcacaatt tccgcgctta tttatggcct aaatcatatt 1980
tcgttggcaa gaacgcacta tatctcgtct cccaatctgc atccgcctga gaatggccgt 2040
attttatgat ggccgcccgc atttgcgtaa cgacagtcag ggtttccccg accagcgagc 2100
accgggtttt gggttcctca agggttcaaa agattaaata ggcccacgca cacgacacac 2160
acacgcacac acacgcactc gcacactgta aatatgcggg ggaagctctt taaaaccatt 2220
tcaactcaat gactgtgccc aaaactcatg tttggtcctt gaaaagtcca gtcagatggg 2280
taaacattct caccaggcat taatttgtgt ctccagctca gataagggtt ggtgatgtgg 2340
gtggcgatat tcatcagctg agcccggttc agccatctgt gatatggttc gccttcggcg 2400
tgattccatg tcccgtcggc ggcagtggaa cagaatgaag cagaaatgtg gccctgacac 2460
ctgtcacaaa tgctaatcgg catttcggtt tcaaggagcc gagcggcatt aatttgctta 2520
tttaatcaaa gtgcatgcca tcggacattc taattatttt cataacggta aggagaggtc 2580
ctcaacctgt ttcacctcct aggactcctt ggaggcgcat ccaaacggat gatgaagctc 2640
agtccaggcc agtcatcctc cattcccgca agattaatcc aggctgatcc ccatttccaa 2700
tgcaaattgc tcgcctgccg ggagctctgc ttctgctcag cattctggtc aatttccata 2760
tgtacgcagt gtcgatctgg gagcctactc acccttctca cccagctcct gttaccctgt 2820
tccttgctcc tttggcgttg tagcactcaa gtcgcagcct aagcagtctg gaattgcttt 2880
taaccgtggc aatattgcga aaatcactct gctatgcagc gtcacgtcgc gattttcttg 2940
tagtgttccg ttccgaagcc cctttccccc atcagctcca tcagctccgc tagctgctcc 3000
acatccttta cgtcccctcg ccgatccttc tgacaccctt ccttgcattt tcatgacttg 3060
tttcctgttt cctgcgtttc cgtttcagac gtgctggtga tgttgtttgt cttgctgcaa 3120
gtttgttttc ggtactgccg gaaccgctgt tctctcctgg cctccaacac cacccccccc 3180
cctatccact gcccccggca tcctgacgta tcctggcgta ccctggcata tccatcttgg 3240
ccgcccccgc actaacttat taacaacata ataacaatgg ctggcaaaca agttgagcgc 3300
tcaggcacac acatgccaat ttgttcagca agcactcggc gctgtctcga gccccagaaa 3360
ggcccagaaa aggagctggc attcagtttc caaggattcc gcgcccccat cctcatccgc 3420
atccgcatcc ttgttcgcat tcccattcct ttccatctgt tgcacggggc gtcgactcgt 3480
gggtcgtaaa tctgaaaatg tcatccctgg ctgggctgag gcaccaattt atggccatca 3540
tcacgtcaag gtgcagcgat gtcctgacac gtccttcctg gctccctgtc ccacctggct 3600
ccttccggag tgggcgtgaa cgcgcccctg gatttatgat cgtaattatg catttgtggc 3660
cagtggcatt tgcattggac tcctgccagt ggcaaccaat cgccacccac tgccattcga 3720
tttcatcatc ttgcttgcaa gagtaatgtt ggcattgtcc attacttatg taaatcgaaa 3780
ccgaaggata agctcctgta agtccatatt ggttatagtt tctgcatatt cgatctacca 3840
agaatttcaa cgaaattcac aagcgaggca gcaataacaa gctcaaagtt agtcggatta 3900
tggggagcta ctaaaaatac atagttgttg acttttggct cagaattaga ggtaataaaa 3960
cacaaaccga tggtcttttt ttatggtcct tttcttttcg caagaaattc ctcctcgaaa 4020
acgaacatct tagcaactaa tattagcaac tttgtatgtt gcccatacca aaaagagtgt 4080
gaaaggaatg ttttctgttt gatgaacact ttaaggctag gacataatct cagataaact 4140
gttgtctaat taacttaatt tgatatccaa ttaaagtaac tatgtgtgct ggtaagacaa 4200
agaaaaggtt tctttgtggc gcttgaaaag gacatttgtg tgcggaaaag ggtcaaattg 4260
acatccttag accaaagttt atggctttgg gctttagaat tttctctatt ccactgctcg 4320
acctggttgc actttttgta ccccaggcgg gacggccttg aaccctgacc tgtggatggg 4380
ccacccattg ggcgcgtcca actcgtccaa gtcgctgaca gctcggcaaa ttaatgatta 4440
attgaatttt ccttggctct gttttgggtt ttttttttt tgttgtccat tctattattt 4500
cttgttttct tggcaacaca aatacaaatc agcgccacac agacgtaatc attaattagg 4560
cgcgcaggga gctaagtgtc tcaagaaatg tgggcgcagt gcaattccag ctagtttttc 4620
aagggccaaa ggcactcgta atgagcaaat ttatgcacaa acacacagtt tctacagttt 4680

ctgatggagc ttggcaaaca ttaaggggcc aaatggcata agtagaaaaa cgaattgaaa 4740
taagaactaa ccacgatgct ctcgttttc ttatttccta tccccccatg tgggcggtac 4800
atgttattga cgcccaaagt gccgccccca tagacccatg ccaccgtgcc acgcgcccac 4860
tctttttctt ttgctggctc caaagtttga acggcatctg cagagaaacg ggcttatttg 4920
actgagaagg aggtaggagg tcggcgaagg acgcatttat ttgaacttat gtacacaagt 4980
gcacatatat acatatatat atatacatac atactgcgct gtgtatgcct gccatattca 5040
tgttcatgtt gttttcgcct tttaatgaga gttaactgag cgggggaact ggaaactccg 5100
cggccaacga tgaatatgaa tgcgccactc aacttaatcc cactcaatca ttatgtggca 5160
atccgcccag ttctttttga ccatgggtct gcaattgtcc tctcggctcg tactcgtaca 5220
caagccccaa cattatttgc accgaacgta tctctgtacc gatctgaatt cgtattgcat 5280
gtcgaactcg acactttatg gtgctccgaa attgttttgt aagcgaaata atggcaagtt 5340
atgacaccca gtgttacaca caattgacat ggttattgat tttcaattat ataaatattt 5400
gcaactaagg aaaaaagcct ctgctgccac tgctgccact gatgtcctgt ggaaaaagca 5460
tttaaatacc gtgttcatcg gcaatctgtg tgcgcttctg tgtattttca atttgcatac 5520
acctgcccat gaagaggcga aagtgatgtg ttaaatgaaa atatgctcct aataaacgtt 5580
tttcatcggt ttcccgtata attgttttct ggctttcaac actacactgc acgaaattta 5640
aggctaacgc gcgagtgggg agcgtgtggc atgaaatcaa tttcattcgt tggttttaaa 5700
gccccgtgtc ccattgattt cgaggcaaaa gcttatattt gtcaagatca taatcaaagc 5760
tgtcggccat taaataaaat gctcaacgag agaaaagaaa cccattaaaa actgagcgaa 5820
ttaaagcgat ataaagcacc attttttgcg tgtgagtatt atttgcgtgg cgaaatgaaa 5880
tatactcctg agctgaatta atgagacccg gtaggcgggt tgaagtgggt gggctttgcg 5940
gacaggcatc ttaagggcgc aaggacatcg ggaacatgga acacggaaca cggaacacgg 6000
agcatggagc atgggcacaa agggaataca aatgaaaggt gatttttcag ctccctttgg 6060
ccttttgagg atttccatgg ggaaaccaca aagtgcggtt gacagggcct aatgtgaaat 6120
taacatggct tatatgcaac gcatactgcc ctgttgcatt catgtcagcc acacacacac 6180
acacacacat ccatgcaccc aaacaccgac agctcaaccg tccatctttg tgtagcgggt 6240
caagttgcca tcaagaacac ggttgtgcca ctgcctctgt tcatcctgtt taacagcccg 6300
gtattggatc ccgaatacca gccaccaccc tccgctccct acaacttgcc gcaaatttcc 6360
ttggacccaa tttattttgc aatttatcgc tggataaaaa gtgactgcga aaagcatttc 6420
aaagcggagg cgcaggacat tgaaaaaaca gtgccatgac aacaattacg gcgtgtttaa 6480
gtagcccgtc tgggtgaaag tcacgccgtg aactgccact ttctcccctt cgttggatgg 6540
cataaaagat agattgttta tagttggtat caaaaggaaa aaggcctggc caatgtcgca 6600
gattgaattg agtgctctgg agggaatcgg ggtcaccaaa tgctggcctc agagtccaga 6660
acgccagctc taagacttgt attattatgg tattaaatgg tcagggacac cagagtacaa 6720
ctttaaagta tcttggtata aagtaaagta tgcagtttta atcctaaact tggctcaaat 6780
gaccttatat tccaccattt ttaagtaaca gtccagtaga tttcaatgca tcacaattag 6840
aaatgaattc agccctgaag gctgaaaaat aaacccaatt aataatttaa agcttcggct 6900
gaacgattag catgcattat taacactcac aaatatatgc cataatatgt acttttatgt 6960
agagcactct caaattgtgt cacattcagc acattccttt cagttcgttt aattaaagtg 7020
tccaataaat taacgctcga caagaaggac aaacttaaat gcacttcaaa tggcgcagaa 7080
cccatcaaga ctttccatca atttgcaaaa gtgttggcca aaagtgaggc tccccacctg 7140
ggaactttgt ttccgtgaac tttgcaccgg cacgcggtga acgcgaaatg cattattcac 7200
ttttatgtct ttttcctcaa ctactgtccg ggacaaaaaa atcagtttgc aactgaagaa 7260
tcttacctcg actcgaactc caattgcgat gtccttgttg cccagctctc agctccttga 7320
tgatgcagcg cagcggcagt tgggctctcc gtaaatcgag cgacataaag cggtcaccca 7380
acattggcca ttcgatggga gtgtgacaaa gctcccagct aaacttcccc ggccatgtca 7440
tgcatgcgaa atggagttct cgagaagcag ctagtgaatc tttaatggat tcatcgattg 7500
cgaagcgaag gaagctgcaa aactttcggt ttttaagcca cgctcggttt ggtagttgga 7560
cactcgataa atttacccaa tcaccagaaa aatatgcaac agcttttggg cttcattcta 7620
ctctagtttt ccaattaaaa tgcaaagttc tagagtttcc aaaactcaat gctcctggac 7680
ccccatcctc gccggcgtcg tcatccttgg gaatggagat ctcactcgat ttggccatag 7740
agtgtgtagg acattttaat ttgttgtgtt ttacaatttg cccccataca aattcatgtc 7800
gattgaggct ttttcgcctg agtccgctca gtttatggca cttttataac tcacatgtat 7860
atatgtatat atgtatgtac agacaggcga ggacccccgg ggcggtaaaa agttgctctt 7920
ttttttcgat ggccttaact tggcgtatca atatgtgggg gggaaaacaa tgggaagcaa 7980
ttttttcaaa taatttcaca ggcgaaatca aaaggaagca gcccgggaag aacaagaacg 8040
cgcggcacat cgaccatctg atacctcttg tctatgctgg tgcggtgtcc attgtgggcg 8100
ccacaagtgg agggtagtga tctgttgata ccatcgacat tcatccgaga aaaaagcgta 8160
tccttgtgcc ttttcaatct ggccatcttg ccatccggcc atccaagcat cggctgccat 8220
cgagcgtcga aaaggccaaa cttttgcaca accgtttttt attgttgact caaatttgga 8280

attattttta tttgagcacc gtggagggac acacgcactg agtccataaa aatccacttt 8340
cagttatgtg gcagtgggtg taagtggaaa tgccgaaaga tggacttatt gtcattgagt 8400
atcatcgcat cgaatggggc gctaaatcaa aaacaacgac cagagggcaa tccggaattg 8460
gattggtttc gaaatctgga attcgcggtg gggggaaatt gaaaatgggc catattttgg 8520
caatcgattt gcatgcgccg cttaaagtca tttccaattc gataatcata atttatgaag 8580
ctgttttgta ttgcaatttt cgcatataaa ttacattaaa tatttacgtc agccgtgaaa 8640
aatacataaa tagtgcgtct ctggtgtgtc cgccgattcc gcgcgagtcc tttgtatccg 8700
catatcctgt tcaaacaata ctgcgctaga gtcgaaatcc aaatatttgc gcagcaaatg 8760
aaaagcattt aaatccggtc gcagaggtta accgtagtag tttttaagcc gcctggctgc 8820
cgcaaccggc aaacttttcg atccactctc gcgactagcg agcaattagc tgcggtcgca 8880
tttcattgtg ctacccgcta tttgcactaa actttacact cattaaacag gcgagaaatt 8940
attgcagcga agtgtccggg attaaagtgt tgcatcccca tcccgctgcc agtcggatga 9000
tgatgaggca tgatgggttt ctgccaggtg gtcgcacttt taattggtat tcaactttct 9060
cctttgaccc tcacaactcg aaacatgctc gtaatttacc aaccaaaata agcgagtttc 9120
ccagttccgt tccaacttcc ttggtcacct caaaagctga aggcgagaag cctgttcgct 9180
aagctaagcc gcactcaaag gttggctctg tgtttgcctt acctttgcag ttgagatgaa 9240
gataaacaga aataaatgaa atgttacatt tgtgtacctt gaaaatattc agaagccatg 9300
tttataaatg aaattataca tttatactat atttaatgc agtgcctggg aaaactctaa 9360
gacattggtg ttgcattaaa aataaagtgc ttaaaagtat gcaacacaga attgatagta 9420
ctgcattagc atgttgattc atctgaatat tgagttacat accccacatg ttgtacttat 9480
cagataatta tttgatttac atgcttaaat catatcagtg ctagcttacc cagaagccct 9540
aaattacttc tatcccaaga cattctctca gtgtacatgc ggtgggtatt cccaaaagtt 9600
taggaaaatt gcccgtggca tttggcttgc cagaaacgga tgcctgtcga agtgcagcta 9660
gttcttttgg gcaaaggcca gttatacttt tgcggttgtt agcttaattt aatgtaactt 9720
tcgacatggt cagatggagt cgaaaggtga gcggtgtatc tggagaagat cgaaaagcaa 9780
ataaccagct ggaaaatgtt tggccacaac cctctgagaa tccacttggc cttacgagtg 9840
gatggatggt tgatgggtgg aaaccggagc acacgcacag cacaattagc agttatgaaa 9900
agttgaaagt tcattttgtg cagcaacaaa acgaaaatgt gatgaaaagt tcgcactgac 9960
cttttggaat ttttcattgt gaacgaagga gctggccgaa aatgattttg ccaaaaagcc 10020
gtgggcgaca ataaacagat gaagtctgag gcatgtttgc ggtccactta attgcaggcc 10080
aagaagctca tgaaataccc tggtagccac aaactggttg ctccatctgc agaccaaacg 10140
aaattatggg attcattcgc agagggaatg tgcaacaaag caactaaaat gtgagatttt 10200
ccatttgaaa gcccacactt tcgttcgatt tgccacaatt tgcagcatgt ataaagttat 10260
atccgtactg ccatgatatt tctgcatggc agaaccattt caagtttgtc actcgaacac 10320
gcacccaagt gcagaatgcc ggcatgtgag gatgcgagga tgcgaaaatg ccaggacacg 10380
cggaggtggc cggtggcttt tcgatagcgg aggatggtca agccacgagg ggcaccgaaa 10440
cgggaacccg gctcattgtc attttccgac ccgagttggc tttgagcacc tgagaaccag 10500
tcgacccagt cgacccagtc gacccagtcg accagtctac ctgtctacct gtcgaccaaa 10560
tgtggacaca ggctccattc gagatggctg ctcgtcaaca gagccatcaa attgtcattt 10620
tccgaacaaa cgcgttggtt ttcactcgaa attggccgaa aaataaatgt ccttgtaaat 10680
cactcgactt ttatctttaa aatatgattt aaattcagcg aataaaatac ccaatagctg 10740
ataatttatg cccaatttat tatgaattaa aaaaaaaaaa aagatatgaa gatttgatat 10800
taaaatgccc aaaaatacca tatatattac caattttatt gattaacatt tgctcatttc 10860
agtcataaac aaatcctttg ataaagccaa taaaataata ttttaaaaaa ttttaaacca 10920
aacccaaatc ataatattaa tcttatgcac accaaactaa atgttgaatt ggcaagacga 10980
gaatgctaat agtacataaa tcacaactta tcactttggg aaacaaccat taaaaatgct 11040
atttaaaagc agacaaacag cgtaggatgg ctggtacatt tttactggga ttttgctctc 11100
ctttggcaat ctgtcatgtt tttgaagtag gtcctcgtcc tcgacagcat cgtctgccag 11160
ataagaaatg catatttaat aattttattc tttcactcag caggggccct aggagcttag 11220
cttgttttat gactgcgaca aagtaatcct gtgcaaaagg taggaagtgg gaaatcactg 11280
gttccaaagg ggcagattat cagggataca acagagatgt gctcgactta gctggcatta 11340
tgagaagctc gtgtggaaaa attggcccca accaacacac aacacaaaaa tgggcgaagc 11400
caaagttctc gatgtcgatg ttgtttccct tcaatcagcg attcaattta cgagtgccaa 11460
aattcgacac agttttcccc ttagccattt tgccacttgc cacttgccac ttttcactgc 11520
cggctgacac gccccatccg gaatgccgcg aaaatgggcg tgcgaatatt ttgcctttcg 11580
ctgaattcat tgttggccca tggcccacgt tgcggtccgg tttctatttt cagtatatcc 11640
ataaaattaa catttttgtt gcggactctc aaaatcactg agctatggtc ctcgtggtcc 11700
tccagtttgc atttcccagc gaaaagggaa aggggcaaga ggtaaagtgg caacggggta 11760
gagggtgcgt atatggcggc gccacgcgac ggagctaaat ttaattttgt taatgagctc 11820
ggcccctttc cgtttcatca tcccgccaat gggggcagcc agccagccag ccagcatggc 11880

gcagccacac tctgccgtcc tattattttg cattcgcaat tctaaaagcc tcaacccgtt 11940
tgtgaggtgg tccgtggttc cagaatccac ttgccaagtg gagcgagtgt gtgtgcgtgt 12000
gttttggcct ttattgcgct cttggtgccc gccaggatta gactgtggat aggcatcggg 12060
attgggattg ggttggggat ggggaatggg gaatgggaat gggaataggg gattgccatt 12120
gggatggcga ttagaacgac aatggaccat tgccggaaac acttgaaaac tgtttcgtgt 12180
tggctgggga ctttgattaa aagcggcacc attgctatca tagcggcaag ccgaaaagca 12240
gaaaacaaaa acaaaacaaa aaaaatgagg gaaaataaac aagttggatc tgacaaagac 12300
attaacattt taattgttcc ggcaacgcaa tggagtcgtg gaccaagcgt aacaattgcc 12360
aatgagtgag ctcgaataaa agtttatggg gtggaggcga agggcagtcg ggaaaatggc 12420
cattaatatt ggagccataa aagtcaaaac ttttcgacgg gtggcaaatg ggaggcgtgg 12480
attagcaata tatcccatgc cgccaggaaa ttgaccctag tcattatttg ccacagagat 12540
gattatgtta ggcaataaac aaacttaagt tttgaatctg ctcgatttgt gaatgaaaac 12600
gccaaaaata tacatatata atctaatgta aatattttag ttatggattt tgaatggctt 12660
ggatttttgt attggttggt aacttctctg tttttccaat tcattgcagc tgcctcggtg 12720
gccgtttcct cctggacact ggtggccatc tcttgcgagc gctactacgc catttgccac 12780
ccgctgaggt cgcgcacctg gcagacgatc aaccacgcca acaagatcat cgccatcatc 12840
tggctgggca gtctggtgtg catgacgccc atcgccgcct ttagccagct gatgccgacc 12900
agccgaccag gtgagtgtgc gccaggataa cactaattca ggaactaacc ccctgtctgg 12960
tcaggacttc gcaagtgccg cgagcagtgg ccggcggata gcctcaacta cgagcgggca 13020
tacaacctgt tcctggacct ggccctgctg gtccttcctc tgctggctct gagcttcacc 13080
tatcttttta tcacccgcac cctgtacgtg agcatgcgca acgagcgagc catgaacttt 13140
ggcagcagtg ggccggaggt caccacctcc tcctctgccg ctgtcgcaga agccggtagc 13200
cagcgacgcg ccaacggaag ccattgtcag tccctcgaca cgattgtgcc acaccagcac 13260
aatcctcacc agcagcacca ccaccattcc cagtactact atggtaaggg gtttatgctt 13320
attttttgag atattcgtag atttttttg tcagccagct gggtcaaagg gaattctggg 13380
ttcagagggt tcagtggcgt agcattaata gttgcgactg ttgccggggg ttggcagggg 13440
gcgtcattcg ggattagctc aactactgag cttatccgcg gacccaaagc ccccgaactg 13500
cgtcacacac aatttgcacg gcaaagtttc ccacgcacga aaagattcgt gtgtctacga 13560
ctcctgccta aatcgcggaa gtgaattgaa tggctggctg gggggatgggg aaaggggggct 13620
aaagcctgga ggagcactct gtccaagttt ttcaagcgtg tgtacgaaga gacaatagca 13680
agcgaaagct ctcccggcgt ttttactttc atttgcattc gcatcgccgg ctttaagagc 13740
accacccatg gctagatggc cccttctccc atctacggcg tcctccatcc tctccagctg 13800
cggatgctgt gaataatccc cagattatgg gcacaaaaat tctaggcagc ccaaggcaaa 13860
gccggcgcaa atcaattctg ctaattcgcc tttaatgccg ccttaaagcg tgctcccaat 13920
atatacatat ctctatgtaa atttagaata tattgtagta tatgctcgtc tgcagcagac 13980
atgcttgact gtaaatatag aacagtccca gtggagaggc ctgctgaaga ggaacctcct 14040
cggtgctcca ccatggatgc gagttgataa cgtgacgtgt atgataaatg cacacccaca 14100
aagcgaagtc aaccgtagag tccccgtgga aatcacagtg agcctaccga gcagcgaact 14160
tactttaatc ttcgagaata ttactctgcc tctggtccaa ttaatgttga gaagaggagc 14220
tcctgccacg gaaagatgga agcactgagg caatggagca gtggagcagt ggagcagtga 14280
gtccctgggc gactcatttg catgaaaggg ttggcgcctc caggtggtaa gtcctttatg 14340
cgaaatggaa agtgtgcatt agcaccttca cttgcgccgc actgaattat cgcagcactc 14400
gagaattcat cgagcgattc atttcctagg caactgaccc cgcttgaccc atggaccaac 14460
acggacatgc cagtccaccc gctgctgaaa aaggtcaatt taaatgccgt aaattatgca 14520
cgagacgagc atttaagcgc tcgttaaagg ccgttatccc attgccagga ttctccactg 14580
gagcccgtcg gacgcgtggc actcattttc acacatactc gcactccggt cgcctttctg 14640
tgtgacaata aatttgatta aatttttaac tttttcgttc cgtttaaggc catctcgttt 14700
gacataaccc cgtcataaat ttccaagcca aagagcattt atgtgtgttg gctggatcgc 14760
aaatgtccca tcgccgaaca agctgaacat tttcatggaa attgccatat tattagcagc 14820
ttcacatggc ccccactatt cgtgccaaga atccgcatgc tatccatccc caaactgcgc 14880
gcttgatggc aatttaaatt atgctcagct gaactactcg cccaaccgca aacggcaata 14940
aaattgcgag gtatatggtc tcattccgca cttaagtatc catttccccc ttacatgtat 15000
cccaaatgat tcccatcaca tgctggtagc tacaaatata gtatcctttc catgtaacat 15060
ctacaatgta gcacaaattt tgtgatatat ggatataatc ccttcattgc atcattaata 15120
tctacgtgtc cttctcttga aactcagcca aatgccattc ctgctggcgg tgaggaaggt 15180
gagcggagac acctccaccc tgcactgatc ctcatccgcc agctggacgc agtggtcctg 15240
cagtagtgtg gtcagtgcag ccttcgtctg gagcaggcca agtcgcattc ccggacacat 15300
tcgtggtcca gctccgaagc ccagaaaccg acagcccatt ggtggctgat ttagaaagcg 15360
ttctggacga aactggttgg gtgccggata caatgcggga tctctgaaat cgaaaaagac 15420
tgtagttcgt agacaatctg aagtctcact gaagtgtcac tcacaaatga atggcctgaa 15480

ccggcaacac cagcacggtg cccaaatgca ccttaagttc gctgctcgct cctgatttct 15540
gatcaggaag ggtgaatgtc ttggtgcagc gcttctgcag agcctgcatg gccggatgaa 15600
gacgcagtgc ctccaagagc gcagcttcgc tatagcgaag ttcacccaga gccactggat 15660
caatcagatt gccagcgtgg cgctgggcca cttcatccaa ttcgatgtgt aaccgtcgct 15720
gtgcatcctc attgagggcc aattcgtaga gggcgaaggc gagtagcatc gccgaggttt 15780
cgtagccctc cagtaggaga gtggtggcat gtccagctag ctcctccttt cccaagcctc 15840
ttttgctttc cgccagccat tgcaagaggt tgtcgccacc acttcgggct tccaccaatt 15900
ccctgaacca gtgctgcagg ggcaggggta cgtatctgta taagtggata gtggatcaaa 15960
ggttacacat tggcgagttc cacattacca accggtggcc aatgagtcta cccaatcgag 16020
gcgtgtgcaa cagtgacatg gtctccagca aactccaaac gcttggctga aagagcggag 16080
cgagccattc cagccagcga cttggttcat gagccaccct catgtgaatc cccaggcaat 16140
gtgcatccag accgaagatc gcggaggcaa ccacttgtag cgtataacgc gtggccaagt 16200
ccttggcctc gaagcgtcca aggggcacct gatcccggag caactgacat gctctggcca 16260
catgcggcaa ggtctgacgc actcggctgg gtgtgaagag cggcaccagg tcggcccgca 16320
agactcgcca gcggtctccg ttggcgatga agggattgtg gctagccagg acatcgcgct 16380
ggtggtccac ataaacggcg ttgtccgcac agtcggcaaa gttacggccc accaggatct 16440
ggcgcagcaa ctcctggtca cgcaccagaa tggcgggttc gttaaacagg cgatagaaac 16500
ccacatattt caggcccgga ttccgactgt caagaaaagt tcattgcacg tggccttgaa 16560
ggactcttaa gattagtgta tcttacgtat agatttcctg gtaaatctct ccatatgagc 16620
ggcggcccag ggcaaagtcc aacatattcc caaccaatgg ccagccaaaa ggtgcctcca 16680
gtcccaggcg acgccagtga ttgccctgcc atttccaaaa caggagcgtg gcggccaata 16740
ggataaatag caccagtggc agcattatcg atgagcatct gaaaagcttg aaaggcacaa 16800
tcccactcaa cttgcattcc cagttgagtt gaatatctca aagcagctga gagtgaacca 16860
cgttgcgtga ttccatgatt gcgcctcatt agccacgtaa ttgctgtata tcaatcaact 16920
gacgggtatg tttcaatcac gcaagattcg agacagctag cattgggctt tgaattaatt 16980
tctataacca ctatgactca atatgttttc ggataataat cttagtattt caacttatat 17040
gcattgtaca catccattca gttttaattt atttagtgca tttcttctgt tcaaaatttt 17100
gttttttttt tctgtgctta acattcaatt tgacaatcgt atccatttgc gagtcaaatc 17160
agatgtgttt tgtggatacc ctgtaacttt ccatcttgca gattacggcc actgtggcag 17220
caagcggcga ttgatcagcg gaggaggacc ctgcgaagga aggaggcatc tatactgcat 17280
gcggagcgca tccgtgaagt ccctgcgcca tcagcagatc aacggaggag gtggaaccct 17340
aagtggaact ggagctggaa atggcgagtg ctgcagccgg gtgcacagga tgcgccaaca 17400
gatgcagctg cagcagcaag gctacgtgag cgacaatgaa tcccggcgca agtcattgtc 17460
gcagcccagt ctgcgtatca cggaagcggg actgcggagg tgagtgcttc agggtattct 17520
catccacata tgaccgaccg attcgtgtgt cgattttcgc tttcacggta ccttgcagat 17580
ccaacgaaac caagagtctg gagagcaaga agcgtgtggt caagatgctg ttcgtcctag 17640
tgctggagtt cttcatctgc tggacgccgc tgtatgtgat caacacgatg accatgctcc 17700
ttgggccgac ggtgtacgag tacgtgggct atacctccat cagtttcctc cagctgctgg 17760
cctactcatc cagctgctgc aatccgatca cctactgctt catgaacgcc agcttccggc 17820
gcgccttcgt ggacaccttt aaggggatgc gggtgtgcga gcgcctgtgc gctccctgct 17880
gcttctggcg gcgacgctcc aagaacgaga cgaatctttc ggtggccggt aactccattg 17940
cgctggccaa ctcggtcatg tccagccaca cgatcctcga gagtccgcga ctctgagcca 18000
gccgccgccg acggcagatg gccgagtcgg agtcacttta ggttagcaat gggtacggat 18060
cgggagcgct aatatgtgat gatgcaatgt aactgcaata gaaactgtgg ttgtcaacga 18120
cattccgccg cgttctgaat gtgtgtggaa gtggccagat ggataggtgg gtggatggag 18180
gaggtgccat ggcaattgcc cctccaaatt gctcttcatc ttaccgataa gtattttaat 18240
atatgcgtat gtgtgtatgc ttagagttgc gctaaagcgc tttattccgt gtgtgttatt 18300
cgactttgac ctcgcgttcc atgtgtttac accttaaggt gggtgcaacc agtcatgcca 18360
atctaaattg agatcgccac acgctatggg gaatggaaag tgctcagctg gattaagatt 18420
tttaaacaaa caataggcta agaaatcagc acaatggttt cacgacatgg tcacagtatg 18480
tcgtgagcca acaaaatgac ttccaaatgg ggggcgctcc ataaaattgt tttaatttat 18540
ggaaaatggt atatgtgttg gaaggatttg ttggttcaaa tctggagttg ggtcaaaagt 18600
tggtaaacaa ctcatttgtc acacatgtcg tatgagtgat atttacgcaa gctatccgac 18660
cattgtgctg cactttctaa ttttatttcc acactttacg ttccaaagat caatcccaaa 18720
agtaaaaggt attctgatca cggtgcagtg aaaggtttca atgaatcaag aagcatttga 18780
cagcaagctg attttacagg gtacacaggg aagtcccact ctcaaaagga ttcccatcat 18840
tcaaaaatgg ttgtcagacg gagcttatgg agtggcagac actccgaggt ccaaaactaa 18900
attaagaaac ccaatcgggt gccaaaatcc aattttaagt agacgaaagt agctggaaat 18960

gccaatcgga tatattcggg cgggattagc attgtg 18996

<210> 62
<211> 2250
<212> DNA
<213> Drosophila

<400> 62
atgttgccgc gcctgtgcgc cgacgcttgc cgtcagtgct tcgcgaagat tgcgaggcgc 60
gacacacatc gcggaacccg aactccatac ggctgtgccg acactcaatc ccgacctaag 120
cccaattttt tgttgcgtga agttgacgaa gtgtgttgta ctgctgcaag tgcaagtccc 180
cgattgctag tgctattccg ggatcacaaa cgagcatcct tcttcggcct cacgattgac 240
gccttctacc actacctgcg ccaggcgctg ccgctggcca aggaggctgc gatccacctg 300
aatgccagca acgagatctc cgctgtcggt gacggggtca ccatcacagg gacacccggt 360
gatcttctca actacagcgg gctggaattg gatctcggac tggatttgga tcttaatctg 420
gacatggacc tagccactac tcccagcagc agcaccttgg cgccagctgt aaccgtcagg 480
acaccgggga acaggagcgt ggtgagggtc agtgccgatg tgcccatctg ggtggtgccg 540
tgctattcgg ccatccttct gtgcgccgtg gtcgggaacc tgctggtggt gctaacgctc 600
gtccagaatc gtcgcatgcg caccatcacg aacgtttttc tgctcaacct ggccatctcg 660
gacatcctgc tcggcgtatt ctgcatgccg gtgaccctgg tcggcactct gctaaggcat 720
ttcatattcg gcgagctgct ctgcaagctc atccagtttg ctcaggtccc tatccccgtg 780
accgagcccc ataatcgctt gtcaaatggc cacgaaccag cacgacctcg agcggattcg 840
ccagcagatc gtgttggcca acattcagga gctgataaag aagatgacac gtcgctgctt 900
cgacgtatgt atcgctatgc cggaaatgga gttgcgctcc acggagcgcg actgcctggc 960
caactgtatg gatcgattca tggactcggt tcagctgcct cggtggccgt ttcctcctgg 1020
acactggtgg ccatctcttg cgagcgctac tacgccattt gccacccgct gaggtcgcgc 1080
acctggcaga cgatcaacca cgccaacaag atcatcgcca tcatctggct gggcagtctg 1140
gtgtgcatga cgcccatcgc cgcctttagc cagctgatgc cgaccagccg accaggactt 1200
cgcaagtgcc gcgagcagtg gccggcggat agcctcaact acgagcgggc atacaacctg 1260
ttcctggacc tggccctgct ggtccttcct ctgctggctc tgagcttcac ctatcttttt 1320
atcacccgca ccctgtacgt gagcatgcgc aacgagcgag ccatgaactt tggcagcagt 1380
gggccggagg tcaccacctc ctcctctgcc gctgtcgcag aagccggtag ccagcgacgc 1440
gccaacggaa gccattgtca gtccctcgac acgattgtgc cacaccagca caatcctcac 1500
cagcagcacc accaccattc ccagtactac tatgattacg gccactgtgg cagcaagcgg 1560
cgattgatca gcggaggagg accctgcgaa ggaaggaggc atctatactg catgcggagc 1620
gcatccgtga agtccctgcg ccatcagcag atcaacggag gaggtggaac cctaagtgga 1680
actggagctg gaaatggcga gtgctgcagc cgggtgcaca ggatgcgcca acagatgcag 1740
ctgcagcagc aaggctacgt gagcgacaat gaatcccggc gcaagtcatt gtcgcagccc 1800
agtctgcgta tcacggaagc gggactgcgg agatccaacg aaaccaagag tctggagagc 1860
aagaagcgtg tggtcaagat gctgttcgtc ctagtgctgg agttcttcat ctgctggacg 1920
ccgctgtatg tgatcaacac gatgaccatg ctccttgggc cgacggtgta cgagtacgtg 1980
ggctatacct ccatcagttt cctccagctg ctggcctact catccagctg ctgcaatccg 2040
atcacctact gcttcatgaa cgccagcttc cggcgcgcct tcgtggacac ctttaagggg 2100
atgcgggtgt gcgagcgcct gtgcgctccc tgctgcttct ggcggcgacg ctccaagaac 2160
gagacgaatc tttcggtggc cggtaactcc attgcgctgg ccaactcggt catgtccagc 2220
cacacgatcc tcgagagtcc gcgactctga 2250

<210> 63
<211> 749
<212> PRT
<213> Drosophila

<400> 63
Met Leu Pro Arg Leu Cys Ala Asp Ala Cys Arg Gln Cys Phe Ala Lys
1 5 10 15
Ile Ala Arg Arg Asp Thr His Arg Gly Thr Arg Thr Pro Tyr Gly Cys
20 25 30
Ala Asp Thr Gln Ser Arg Pro Lys Pro Asn Phe Leu Leu Arg Glu Val

```
        35                  40                  45
Asp Glu Val Cys Cys Thr Ala Ala Ser Ala Ser Pro Arg Leu Leu Val
    50                  55                  60
Leu Phe Arg Asp His Lys Arg Ala Ser Phe Phe Gly Leu Thr Ile Asp
65                  70                  75                  80
Ala Phe Tyr His Tyr Leu Arg Gln Ala Leu Pro Leu Ala Lys Glu Ala
                85                  90                  95
Ala Ile His Leu Asn Ala Ser Asn Glu Ile Ser Ala Val Gly Asp Gly
            100                 105                 110
Val Thr Ile Thr Gly Thr Pro Gly Asp Leu Leu Asn Tyr Ser Gly Leu
        115                 120                 125
Glu Leu Asp Leu Gly Leu Asp Leu Asp Leu Asn Leu Asp Met Asp Leu
    130                 135                 140
Ala Thr Thr Pro Ser Ser Ser Thr Leu Ala Pro Ala Val Thr Val Arg
145                 150                 155                 160
Thr Pro Gly Asn Arg Ser Val Val Arg Val Ser Ala Asp Val Pro Ile
                165                 170                 175
Trp Val Val Pro Cys Tyr Ser Ala Ile Leu Leu Cys Ala Val Val Gly
            180                 185                 190
Asn Leu Leu Val Val Leu Thr Leu Val Gln Asn Arg Arg Met Arg Thr
        195                 200                 205
Ile Thr Asn Val Phe Leu Leu Asn Leu Ala Ile Ser Asp Ile Leu Leu
    210                 215                 220
Gly Val Phe Cys Met Pro Val Thr Leu Val Gly Thr Leu Leu Arg His
225                 230                 235                 240
Phe Ile Phe Gly Glu Leu Leu Cys Lys Leu Ile Gln Phe Ala Gln Val
                245                 250                 255
Pro Ile Pro Val Thr Glu Pro His Asn Arg Leu Ser Asn Gly His Glu
            260                 265                 270
Pro Ala Arg Pro Arg Ala Asp Ser Pro Ala Asp Arg Val Gly Gln His
        275                 280                 285
Ser Gly Ala Asp Lys Glu Asp Asp Thr Ser Leu Leu Arg Arg Met Tyr
    290                 295                 300
Arg Tyr Ala Gly Asn Gly Val Ala Leu His Gly Ala Arg Leu Pro Gly
305                 310                 315                 320
Gln Leu Tyr Gly Ser Ile His Gly Leu Gly Ser Ala Ala Ser Val Ala
                325                 330                 335
Val Ser Ser Trp Thr Leu Val Ala Ile Ser Cys Glu Arg Tyr Tyr Ala
            340                 345                 350
Ile Cys His Pro Leu Arg Ser Arg Thr Trp Gln Thr Ile Asn His Ala
        355                 360                 365
Asn Lys Ile Ile Ala Ile Ile Trp Leu Gly Ser Leu Val Cys Met Thr
    370                 375                 380
Pro Ile Ala Ala Phe Ser Gln Leu Met Pro Thr Ser Arg Pro Gly Leu
385                 390                 395                 400
Arg Lys Cys Arg Glu Gln Trp Pro Ala Asp Ser Leu Asn Tyr Glu Arg
                405                 410                 415
Ala Tyr Asn Leu Phe Leu Asp Leu Ala Leu Leu Val Leu Pro Leu Leu
            420                 425                 430
Ala Leu Ser Phe Thr Tyr Leu Phe Ile Thr Arg Thr Leu Tyr Val Ser
        435                 440                 445
Met Arg Asn Glu Arg Ala Met Asn Phe Gly Ser Ser Gly Pro Glu Val
    450                 455                 460
Thr Thr Ser Ser Ser Ala Ala Val Ala Glu Ala Gly Ser Gln Arg Arg
465                 470                 475                 480
Ala Asn Gly Ser His Cys Gln Ser Leu Asp Thr Ile Val Pro His Gln
                485                 490                 495
His Asn Pro His Gln Gln His His His His Ser Gln Tyr Tyr Tyr Asp
            500                 505                 510
Tyr Gly His Cys Gly Ser Lys Arg Arg Leu Ile Ser Gly Gly Gly Pro
```

515 520 525
Cys Glu Gly Arg Arg His Leu Tyr Cys Met Arg Ser Ala Ser Val Lys
530 535 540
Ser Leu Arg His Gln Gln Ile Asn Gly Gly Gly Gly Thr Leu Ser Gly
545 550 555 560
Thr Gly Ala Gly Asn Gly Glu Cys Cys Ser Arg Val His Arg Met Arg
565 570 575
Gln Gln Met Gln Leu Gln Gln Gln Gly Tyr Val Ser Asp Asn Glu Ser
580 585 590
Arg Arg Lys Ser Leu Ser Gln Pro Ser Leu Arg Ile Thr Glu Ala Gly
595 600 605
Leu Arg Arg Ser Asn Glu Thr Lys Ser Leu Glu Ser Lys Lys Arg Val
610 615 620
Val Lys Met Leu Phe Val Leu Val Leu Glu Phe Phe Ile Cys Trp Thr
625 630 635 640
Pro Leu Tyr Val Ile Asn Thr Met Thr Met Leu Leu Gly Pro Thr Val
645 650 655
Tyr Glu Tyr Val Gly Tyr Thr Ser Ile Ser Phe Leu Gln Leu Leu Ala
660 665 670
Tyr Ser Ser Ser Cys Cys Asn Pro Ile Thr Tyr Cys Phe Met Asn Ala
675 680 685
Ser Phe Arg Arg Ala Phe Val Asp Thr Phe Lys Gly Met Arg Val Cys
690 695 700
Glu Arg Leu Cys Ala Pro Cys Cys Phe Trp Arg Arg Arg Ser Lys Asn
705 710 715 720
Glu Thr Asn Leu Ser Val Ala Gly Asn Ser Ile Ala Leu Ala Asn Ser
725 730 735
Val Met Ser Ser His Thr Ile Leu Glu Ser Pro Arg Leu
740 745

<210> 64
<211> 4289
<212> DNA
<213> Drosophila

<400> 64
agctgtaaag tactactact acatatgtgt acattcaagt ttcgcttacg aatgcaacgt 60
tccattaaag aagccaccaa ctttaagata cacgcatata acaactcctg agaattgtga 120
aaactgatct actttgatga tcaactttca ttgctgctga ttgaataatt gagtaatgag 180
taattggtat gctgataaac aggaaaacca aaatgtcaac agttgataag gaacattaaa 240
acacaaatta agaggacggt aagtccttca ctcggacctt ttcagctgag caccgtaaag 300
tgctgaaggg attgccccca cacccagacc gcgcccactt ccacttccac tttaaaataa 360
tcatcatggc cctctttttt ggccaggctg cggtggcttt gggccttaat tgtcgactca 420
gttgctgtgc tgtgctgtgt actcgacaag gcgacgaacc gagagccttg cagctccata 480
atgacaggcg gatggactcg aacactcgga cccggacccg gacccgaacc cggacccagt 540
ccacagccag gacataagcg gcttgtgctc cactcaactg tcgcctttcg gagctgctcc 600
tagcttgcaa ttaaaatacc cacttgggcc ctataaattg tgaacgtagg aaggcaggcg 660
gaaaaaaaaa aaaatattcc agcactttgg tgtatggagt ggggtccact taattcccag 720
gtactcaccc aaaaggattc caactataaa gacatacagc gagagtgctt aagttttact 780
ctgtttattt gcagcaacct taggaagttc aattgaaaca ccaatagact agcgtctgag 840
ttattttacg agtacgatca taatcacaag taggtaaaaa aaatccatat tcgccttact 900
taaaacaaaa aaaaaaacca cacgcagatc gcttatcata catatttaca ccgcatctcg 960
ttgatcctga atgggtcctt ggaggcgaat cctctacagc ttagagctga ggactgttga 1020
cggcggcggc ggcactattg gtggtcacca ccgcattgag aaactcgtgg cgtgagcgac 1080
tccgccaggt gcccataccc atggctaggc cgggattgag actaatgttg gtgttggcac 1140
tcgcataggc gcccgggcct gcgcccgcct ggctggcgga gagtcctcca ccagcagcac 1200
caccgacgcc tccgctggca cctgctccac gacgccaggg cagacccttg aaggtgtcga 1260
caaaggcgcg ccggaagctg gcgttcatga agcagtaggt gatcggattg cagcagctgg 1320
atgagtaggc cagcagctgg aggaaactga tggccgtgta gtcgacatac tcgtacacca 1380

ccggtccgat cagcatgacc atcgtgttga tcacgtacag cggagtccag cagatgaaaa 1440
actccagcac caggacgaac agcatcttga ccacacgctt cttgctctcc agggtcttgg 1500
cctcgttgga cctggatagt ggattgtgga ttgtggagaa atgggaatag ggaataggat 1560
tggcagaccg gttatcggac acgaatccta cctgcgaagt gccgcatcgt ggacgcgaat 1620
gctgggcgag gaggtcttgg ccagcgtcac cgtggtcgtc gtggtggtgg tgatcacagt 1680
tggagccgtt ggtctcgttg tcaaggtggt cgttgccata ttggttgatc ctccgcctgc 1740
ggatccctct gaatttccat tattattgtt actattttct atgtgaaaaa aaaaaaacat 1800
gtgaacacaa gccaaatata atcaaacaaa ttaacgttat tagcagcaac aaataaaaca 1860
gttgccataa ctaatctaaa ctaatacata gacgtaaaat aattaaattt tcaatttggt 1920
tttaagagtc tcaaaaataa tatagcatat tatgttaagt ggatttgtac cttcttatcg 1980
agcatttcac caacaatcta tatctttaaa gtattatgcc atattacatg tcatgtaaca 2040
ttaaataact tatgagttat gatatttaga tgcaaaataa tacttgaatg cattcataaa 2100
tgctcttaaa gttgtttgca tattcatggc ttttaaggat tgaaaaagga ttagtttcgt 2160
gcagttacca accattatag actgcggtgg cggtcaggac caggatgcag ttactactgc 2220
tgctggtgcc cggattcggt gcgcttccgc cggccgttgt agcggaaaca ggcagcgatt 2280
gctgcaggat gcgtccgctg tccttggcca tgcctacgta cagggtacgc gtgatgagga 2340
tgtaggccac gcagaggacg agaagcggca ggacgagcag caggaagtcc agcaggatgt 2400
tgtagaagag ctcgtatccc tggtcgggcc aaaactcacg gcacttgcag tagcctgtaa 2460
cgtaagttgg tataacaacc agacggtgtc tgctggccag gttgctatgc actcacccgg 2520
tcgactggtg ggtatcaatt gactaaagac cgctatgggc gtcatgcaga ggatgccgcc 2580
cagccagatg aagccgatga tcttgtaggc gtgactgatt gtctgccagg atcgcgagcg 2640
cagtggatgg catatcgcgt agtagcgctc acaggatatg gccaccaagg tccaggacga 2700
aacggccacg gaggcggctg aaataaataa atgaaagcgg attaacttac aacaaaccga 2760
acacaatgga cacacccagc agctcctgtc gagcgccgtg tcaaatcgct ttcattagaa 2820
ggtccttcgc ccgcgttcgt tgcattttaa ttacacaatt tgcatattgg gaccagacga 2880
accgtaccaa ccagacggac ggagacatcg ctcttcgcac ttctgcctgt tcttgtagta 2940
tgacaaacac gatgcacgat gcaatttgca gaaagtgcaa ttaagcaatt ccccagccac 3000
tcctttctt gttgcccggt ctcctttca tgtaagcctg ttagatgatt tcccattgta 3060
aagcaatttg tacccacacc ctgcccatcc ctcatattct tttgggctac cactatcgat 3120
ggggctacac aagtgcaagt caatagggtg gactgctgta ggtgaaaaac tcgacaccga 3180
aagaaaaaga cccgattcga aaaattggtc aattttagcg caattaaatg agaaataatt 3240
tctagttttt caatcgattg ttttatagca ccaaatcata tatcttaacc gtaaaggcat 3300
caaataaata tgataaactg atcaaatgaa agcataattt gcttgtatgt atgcgtgttt 3360
ttggaattta tagcttacca aaacacttgt gaaatttgca aatatacata tttgttcata 3420
cacaatttag acttcaatgc aataagctgt aaggtctatg gtttcaatta ttaatactcg 3480
gaaattgttg tgctagcaga tacatttgtg gtttgttaga tatacatatt tgcttataaa 3540
gatatatatg tatgtacgta ttgcattcaa aagcaataag cagtaaggta aattggatca 3600
agggacgagt atttaagaca gcctgtatat gttcccttta ctttaatgcg acacctcatt 3660
caggcaatta aggtccaccc taatgtgctt atcccttcag ccatctgcag ctgcattgta 3720
ttatattgga atggtatttt ctcgcaaaat gatatgcaag ttttagattg atacgtacaa 3780
ttgcagtgat tactacttta gttacgaatg atgataggtg atgtagtaga tatatatttg 3840
ctaccaaaga tagttacacc ttagaatcct aaaagaaata agtgttttgt agtcgcagtt 3900
cgtttctcta gttcaacaca tccagccaac caaggtccac cctaatgtga ctgcacttca 3960
gctgcactgt attatattac atttagctg cacaattgtt tactttctat tttaattatg 4020
catgcaaact aatatgagca gtaaagatgg attacctcca cctgttccgt tggagtgcat 4080
cccgtccagg caaccaaggt ccaccctaat ggggtagccc ttgcccttcc aactgcattg 4140
tcacattggt ctcgccctct gtcgcgtttg ttgagttaaa cccctgaaat tgaagtttac 4200
ttttgattga cggcggtgtg tgatgctcta agccagaaat gctagcatgt gtgcgcttaa 4260
acttaaaaca aacagcagca tggcatttc 4289

<210> 65
<211> 1248
<212> DNA
<213> Drosophila

<400> 65
atgcactcca acggaacagg tggagccgcc tccgtggccg tttcgtcctg gaccttggtg 60
gccatatcct gtgagcgcta ctacgcgata tgccatccac tgcgctcgcg atcctggcag 120
acaatcagtc acgcctacaa gatcatcggc ttcatctggc tgggcggcat cctctgcatg 180
acgcccatag cggtctttag tcaattgata cccaccagtc gaccgggcta ctgcaagtgc 240

```
cgtgagtttt ggcccgacca gggatacgag ctcttctaca acatcctgct ggacttcctg 300
ctgctcgtcc tgccgcttct cgtcctctgc gtggcctaca tcctcatcac gcgtaccctg 360
tacgtaggca tggccaagga cagcggacgc atcctgcagc aatcgctgcc tgtttccgct 420
acaacggccg gcggaagcgc accgaatccg ggcaccagca gcagtagtaa ctgcatcctg 480
gtcctgaccg ccaccgcagt ctataatggt tggcggagga tcaaccaata tggcaacgac 540
caccttgaca acgagaccaa cggctccaac tgtgatcacc accaccacga cgaccacggt 600
gacgctggcc aagacctcct cgcccagcat tcgcgtccac gatgcggcac ttcgcaggta 660
ggattcgtgt ccgataaccg gtctgccaat cctattccct attcccattt ctccacaatc 720
cacaatccac tatccaggtc caacgaggcc aagaccctgg agagcaagaa gcgtgtggtc 780
aagatgctgt tcgtcctggt gctggagttt ttcatctgct ggactccgct gtacgtgatc 840
aacacgatgg tcatgctgat cggaccggtg gtgtacgagt atgtcgacta cacggccatc 900
agtttcctcc agctgctggc ctactcatcc agctgctgca atccgatcac ctactgcttc 960
atgaacgcca gcttccggcg cgcctttgtc gacaccttca agggtctgcc ctggcgtcgt 1020
ggagcaggtg ccagcggagg cgtcggtggt gctgctggtg gaggactctc cgccagccag 1080
gcgggcgcag gcccgggcgc ctatgcgagt gccaacacca acattagtct caatcccggc 1140
ctagccatgg gtatgggcac ctggcggagt cgctcacgcc acgagtttct caatgcggtg 1200
gtgaccacca atagtgccgc cgccgccgtc aacagtcctc agctctaa             1248
```

<210> 66
<211> 415
<212> PRT
<213> Drosophila

<400> 66

```
Met His Ser Asn Gly Thr Gly Gly Ala Ala Ser Val Ala Val Ser Ser
 1               5                  10                  15
Trp Thr Leu Val Ala Ile Ser Cys Glu Arg Tyr Tyr Ala Ile Cys His
            20                  25                  30
Pro Leu Arg Ser Arg Ser Trp Gln Thr Ile Ser His Ala Tyr Lys Ile
        35                  40                  45
Ile Gly Phe Ile Trp Leu Gly Gly Ile Leu Cys Met Thr Pro Ile Ala
    50                  55                  60
Val Phe Ser Gln Leu Ile Pro Thr Ser Arg Pro Gly Tyr Cys Lys Cys
65                  70                  75                  80
Arg Glu Phe Trp Pro Asp Gln Gly Tyr Glu Leu Phe Tyr Asn Ile Leu
                85                  90                  95
Leu Asp Phe Leu Leu Leu Val Leu Pro Leu Leu Val Leu Cys Val Ala
            100                 105                 110
Tyr Ile Leu Ile Thr Arg Thr Leu Tyr Val Gly Met Ala Lys Asp Ser
        115                 120                 125
Gly Arg Ile Leu Gln Gln Ser Leu Pro Val Ser Ala Thr Thr Ala Gly
    130                 135                 140
Gly Ser Ala Pro Asn Pro Gly Thr Ser Ser Ser Ser Asn Cys Ile Leu
145                 150                 155                 160
Val Leu Thr Ala Thr Ala Val Tyr Asn Gly Trp Arg Arg Ile Asn Gln
                165                 170                 175
Tyr Gly Asn Asp His Leu Asp Asn Glu Thr Asn Gly Ser Asn Cys Asp
            180                 185                 190
His His His His Asp Asp His Gly Asp Ala Gly Gln Asp Leu Leu Ala
        195                 200                 205
Gln His Ser Arg Pro Arg Cys Gly Thr Ser Gln Val Gly Phe Val Ser
    210                 215                 220
Asp Asn Arg Ser Ala Asn Pro Ile Pro Tyr Ser His Phe Ser Thr Ile
225                 230                 235                 240
His Asn Pro Leu Ser Arg Ser Asn Glu Ala Lys Thr Leu Glu Ser Lys
                245                 250                 255
Lys Arg Val Val Lys Met Leu Phe Val Leu Val Leu Glu Phe Phe Ile
            260                 265                 270
Cys Trp Thr Pro Leu Tyr Val Ile Asn Thr Met Val Met Leu Ile Gly
        275                 280                 285
```

```
Pro Val Val Tyr Glu Tyr Val Asp Tyr Thr Ala Ile Ser Phe Leu Gln
    290                 295                 300
Leu Leu Ala Tyr Ser Ser Ser Cys Cys Asn Pro Ile Thr Tyr Cys Phe
305                 310                 315                 320
Met Asn Ala Ser Phe Arg Arg Ala Phe Val Asp Thr Phe Lys Gly Leu
                325                 330                 335
Pro Trp Arg Arg Gly Ala Gly Ala Ser Gly Gly Val Gly Gly Ala Ala
            340                 345                 350
Gly Gly Gly Leu Ser Ala Ser Gln Ala Gly Ala Gly Pro Gly Ala Tyr
        355                 360                 365
Ala Ser Ala Asn Thr Asn Ile Ser Leu Asn Pro Gly Leu Ala Met Gly
    370                 375                 380
Met Gly Thr Trp Arg Ser Arg Ser Arg His Glu Phe Leu Asn Ala Val
385                 390                 395                 400
Val Thr Thr Asn Ser Ala Ala Ala Ala Val Asn Ser Pro Gln Leu
                405                 410                 415
```

<210> 67
<211> 9249
<212> DNA
<213> Drosophila

<400> 67

```
caaaataaat gtatataatt tcaaaattga catttgctaa tgataagcct ctgcctcatg 60
cccatgctta tgctgaaata tggacaaaat tcgggcacag agcaagagat gcgcagagca 120
gctgccatcc ggaatggaat gagttggaat ttgggcagca gatggcgctg tcgcaaaagg 180
aaggaagcta aatttcactg ggccccaacc aaacggaaat ttccacccac ggctcagaaa 240
ttgaagctgt ttatggtaaa ttaaaccaat tgagtgggat tacgttttgc gcactacaac 300
aacgccacca acctccgcat caactccgca atctccgaca ccccgaagtg tttccaactg 360
ggtttcaatt tcactgtctg gacgagcctg cggctaacga aatttatata cccatacaat 420
gccatcaaag ccacaaggtg tttatgggct taactccggc cctgattccg cctacttggc 480
tgcctcgcct cacagatttg gatttgcatg tgccacgatt tggggcatgc ggctgagaag 540
gcggctgagc aggcggatag aagacaaata gccagcacgt tgcggttcat ttggctgtca 600
caacaaatgc cgaattgatg caaatgtaaa tggattcgat atccaggaac gcgtcgtcgg 660
aatcccaagc agcagtcaag tgatgaatac acggcttaaa agatactgta ttaacttaca 720
gtagcttacc ttcaaagcaa aaattgtaga tttcattctc tagtaacgag ttcatatcca 780
tacctcttta atattggata ctatttatta cccagtttat tacttagttt ttatttcaga 840
atcaaaaaca aatataagca agcttacccc atttaagtac ttggcaattg caaatgtgca 900
cataccactt tcttttccat ccacaaactg gctttcaaac tgctgttggc caactcataa 960
atatggcaga caataaatct agattctttt tcaattgcag ggtttggcaa acgattggca 1020
ataataggag tcatcaaaac gagagcaaat gtcatacagc atgtcgtcct tattccgtgg 1080
cgttagtgcg acatttggag ttggaaacta ggagcccact tagactaaca acgcccgtga 1140
ccattaccat gacgctactg cagaggcttc aggccatgtc ggccaccacg accaggacaa 1200
tactggaggg cagcatcagc agttttggtg gcgggacaaa tgagcctctg gcgagcaaaa 1260
tacccgttct ggaggagtca gcctcacatg ccagatattt gaaattcatt gccgacgggc 1320
tcatcgacga gggactgggc agtgcggtgg gcagtgggag cagcatcgcc gtatccgttg 1380
aagacgtggt cgccggacag gcgcaggaca tccaggcgag cgaaggatcc accgacgacg 1440
ccgacggcag tagccatttg gcattagtct tcgtcaagtg tttcattatt ggtttcatca 1500
tactggccgc catcctgggc aacatgctgg tgattgtgtc ggtcatgcgg caccggaaat 1560
tgcggtgagt ccttgccatt gtgcgttcta gccatatccg ttggtcagga catgtgtgtc 1620
ctcaaatgtc tgggtttcat taatttgtcg ccatgaatca ggtggctttt ttcccttttt 1680
gcgccatggt ttcaatttca ttgtgtgcat ttggcaggaa gcttccaact ataagttact 1740
tagataataa tttaacaaag catgaacttc aacatcaaaa cattagtgag caaatattaa 1800
aagagccctt actttcgctt ttgtttaagg taaattgccc gagctgtaac ctgccacatt 1860
gtccacgaat cgtgggtgtt ttagctgtca aggacatgac ttgtacaccc cacccccccc 1920
gccccttttt ccttttattt ctttgtttgc tcctttttgc cgaacgattt tgtgcgattt 1980
gaccgcgttc ggcttaacga gtagctacct ccttttttttg cggccattaa tttgtcatca 2040
tttatcatat tttccatcat ttttgcgtgt caagcgcaca gaactcatct cgttcgccca 2100
ccgaacactt cagagggggg cgtggttggt tgactaagtt aaatgctaga tacgtgcgta 2160
```

attacattag aaatgttaat aaggggcgag cataaaaata ttgttgccta cttacggggg 2220
cgccattggc atttacttta tgtgcggcca tataatttgt caaatttatg aattgggcgc 2280
caccttcaca tgagatatgg ccacgtgggc gtttgaggga tttccaggta tccaggtagc 2340
aaagtacgta cgtgtttcag tcgcaaattc aaaagcccaa caccgctggc atttaaagaa 2400
cccaacataa aacacttgga gagaggcgac aatggtgtaa cctaaccccg ggacattcaa 2460
tgtcaaggcg cagactatta aagcagcccc gggaaaacgt cttccactgt cattgccgta 2520
ttcctttcgc cttggaacca accaacgcac acctgacatt tgtggctgcc accgagaatg 2580
aggatgggca tggtgatgag catgggggat gaggctcatt tagcctgcca gggacaggca 2640
atttgggacc tgaccctgac tactacgtgt atgccatgtt gccgttttta tctcaactgc 2700
cttgcgccag catctctttc attattgttg ggctcattgt catgtgcgtt ttccgcacca 2760
ttttcacctt ttccgcctcg ctaggcagtt taaatcatca cggcaaatta gacttaatta 2820
cgaaacaaca tgcattccca cttaggtccc atgctaccta aaaaccttgt gaccaagcgg 2880
agtagcaggt atctatttca atgttctgga taagccatct gggcgacagt tgtgcgactt 2940
tttcacgact ggccactcaa cttacattat tcatgtcaca cgcatcgata aagtgggtga 3000
caaattaaag tcagactgca gatcaagtac gaggaagagc cgggaaagtg tttcctcact 3060
tttcctatta attagagctt tccacgttgg ctgcctgccg aaagtgtcga ctcataattg 3120
taatcaaact tgacacccaa caagtttcct cctttcattc ccgagcaccc actatataat 3180
cggcttttgg gtaacgcctc aataacattt tcgtgttgcc tctgatgcgg gtgttttgcc 3240
tttgtacttt ttgggcgctt taagcttgct ccaaaaatgt agacctttcg ggcgtcagt 3300
ttgtggccgt aattaggttt gcgggagatg gtttaatgtc ccagtgaaat atggcaaaat 3360
ggcgagcgat tgttaacgag tttaatagat aaaggtcgga tacacataaa aagaatatcc 3420
atgattgatg cagtttaaga agttcccatc caaagttcct tcattcaaaa gtttatctgg 3480
cacattaatc cgtttagacg gaaagcaaaa ctttcatttg cgatgttgtg gtacgaagag 3540
tcaagttccc gcagatgtta tggttttaaa catgaattcc aaccgcccat tcggttgcgg 3600
ttatatccgt ggaatggcgt tatgtacgcc tgattgtttg gaaatcagtg cggacggccg 3660
ataaacgcat ttgctgctcc acatagcgta tacgccgcgt gcacctaatt gacttaacgt 3720
tttgcggtgc cacgcaacgt tgatgatcgc tggtggaagt tgcggtggcg ttaagctaat 3780
tattaattaa cccattaaac aatgaaatcc actcagagaa gcgccttttc gtacatgaaa 3840
cggcggtcac acgcggcgta tgcgcaatgt gcgcaccaaa gtttccaact caagttggac 3900
agcaatccga agtaacgatt agcgctgcga gctctgagtg acttgaaagc ttaatgcaca 3960
aatttaattt gagaacttgg cactctcaat accactggca ctcaggcttt aggttcattc 4020
tgccattttt ggttgactgt gtggcttccg ccttggttcc tgctaagttt taacctagca 4080
tttactcctc ctccttgcag catcattacc aactactttg tggtctctct ggccgtcgcc 4140
gacatgctgg tggccctctg tgcgatgaca tttaatgctt ccgtcatgat ctcgggcaag 4200
tggatgtttg gttccgtgat gtgcgacatg tggaacagct tcgacgtcta cttctccacc 4260
gccagcatca tgcacctctg ttgcatatcg gtcgacaggt gagtcccggg tttactttgt 4320
cattgtggaa taattgaagt tccgagtgct actgggatgt cttttatacc ttcctttaaa 4380
tgcccatcaa attgaatagt tgcttgacag tttctagtga caatcttttt tcccctttac 4440
ccgtagatac tacgccattg tgcagccact ggactatcca ctaatcatga cacagcgacg 4500
cgtgttcatc atgctattga tggtgtggct atcgccggcg ctcctctcgt tcctgcccat 4560
ctgctcggga tggtacacaa cgaccgagaa ctacaagtat ctcaaatcga atccgcatgt 4620
gagtagctta cttgaagccc ttaattggga caattaaata tttgtacaat agtgtataac 4680
tctattatgg agcccacaga tttgatgagt aggtcgatcc ggtgtatgca cctacattat 4740
ttattcaatt ttcttgcaca tcaaactaat agaaaattct aatttagttg aaacggtatc 4800
gactttgcat gcattttgca tgccttcgat gccacgtggg cgtccatgtg gcagtggcaa 4860
cctttgactt ttccccccgac ccgcctcaac ggaaattgat ttatgtttgt gtggctgaag 4920
cagttttacg cctttcaggc tgccagccga caggtagcat atttattact ccattttatt 4980
ggtttcctcc gcaaatactt ggcaaactcc ttggcagcac ctcctccatc tacaaatata 5040
tacatttata tatatatata tatatatatg ttctctgccc gtcagcagct gctggtactg 5100
attccccccg ccactgcggt tatcattgcg caataaaatg ctgatgtaga aataaatgta 5160
ggtaattcca ttccggtttt ctgcgctccc attcaggctg acttcagcat aggctgtgca 5220
taattttcgg ctctacacct ggaaggacgc ctgaaacgcc tgacaggcct ttgcatacat 5280
tagcatccgc ggtgggcttt aattagtttt acttgtttat gctaataact gtcagcgcga 5340
agccggcaga atgtcctggc cgtaaatctg attagatgtc ttctggtggc cttttgtcat 5400
ggcagcgctg taatgcacag cgtaattgag tttttattgt cgcacttggc attgcagtaa 5460
tcgcattgcc attgcaattg ccgtcgcaat agccgttgct catacgcaac gttggccgcc 5520
gccagccgca gttggcaatt gccgagcatt acgctccatt acggttcgcc actggcatcc 5580
gactgcactg ctccaacgtt cattttgatt tattgatagt cgctaattga gttgcccatc 5640
aaatcgcaac gatttcgatt ccgatttcga tgctcccctc actccatatt tcagatatgc 5700
gagttcaaag tgaacaaggc atacgccata gtcagctcgt cgatgagctt ctggattccc 5760

ggcatcgtaa tgctgtcgat gtactaccgc atttaccagg aggccgaccg acaggagcgt 5820
ctggtgtaca ggtaatagct attcgagtgg cccctccatt tgcccgcttc caattaactc 5880
taattactta ctgtgtgcgc ccagatccaa ggtggccgct ctgctgctgg agaagcatct 5940
gcaaattagc caaattccca agccccggcc gagcattcag gtggagcagt cgaccatctc 6000
gacgatgcgg cgtgagcgga aggccgcccg caccctgggc atcatcatga gcgccttcct 6060
catctgctgg ctgccgttct tcctctggtg cgtatacttg atggctcaaa ttgggcaatt 6120
atcgctagca ctctaaggat caagtgccga gaagttttta attaactcgg gctcatatca 6180
aataatttgc actttccatg ctacaggtgg gcggaaagtg tgtataaaca attaacaata 6240
caagtatgtt tccttgggat tgaaatataa gaaaggcaat gtagacttaa gcagtagtta 6300
ttatgcacta ttcaagtatc tattccacct acttaaaagt aaagaatcta acactttccc 6360
tttctccaat taaaccgcag gtacatcgta tcctcgctgt gcgatagttg catcactccg 6420
cgcctgctcg ttggcatcct gttttggatc ggctacttca actcggccct gaaccccatt 6480
atttatgcat acttcaaccg cgacttcagg gccgccttca agaagaccct caaggtgagt 6540
tgcgggctgg ccatcgagta atttgcatat accttccact tcgacatctt gatgggactt 6600
ttattctatt caatttcgct ttgttcgggc atccaaaact gaattctctg caccacctta 6660
gacttagtac agggcttcgg caactgctgc aatcgcaaag gacagggcca acgccaccca 6720
ttgcacagaa gcggcggcgg ccaccatttc gacaacagta tcgactggca acgcggaggc 6780
gccaccacca gcgccccacc gacgaggctc acgtttagca attattttgg tagctattgc 6840
atagcatcac gggtgcactg agaaaaaaca tgaacaattt atcaaatata ttaagtaact 6900
tggagtaact tagttccagt tacacgacta acaaataaga aaaaagatat ggaaaaaaga 6960
acattgtctc ctcatttttt ctcagtgcag ctcagtgttc cgcggagagc gtagatagcg 7020
tagcgttcat attagttgat tgaaatgctt cttcatccac atctactggt tgctaagcaa 7080
acaaaatagc ttttcctctg cctgaagaac tgaaaactct gcacgttctt tgctctacag 7140
ctctacagct ctgcctgcgc ctgcctctgt ctctctatca gtttcattct cgagggtatc 7200
tttcgggcat ggtgaatggg acagtgaatg gggttgcaa tgcttgggct ttctaatcgc 7260
ttgaacactc tgcttgccac ttaccatcgt atactaccat ttgccatata tagtgccata 7320
taccatatat catttctctt tgcacgcata gctggcgctt actgtgctga ttcgttgcct 7380
atggcccaac aaaagctaag cccgccaaag tatgcaacag caaaacagaa agcataaaac 7440
gtaatattta aattaaaacg gaacctcttc tttgtatgta tgcccagcta caaccagcta 7500
cacctacacc ttctccacat tcatacaagc gtaaatgctg atgatgaaat gttattcttt 7560
ctctaaaaat tctccctttt atgctgcaca cttcgacaca ccagcaacac caacaccaac 7620
acacacacac acactgaggc acatggacgg gtccttgttc gtcccttttt tttatatgaa 7680
tgcgcggcct ttcatttcgt ttagcatttt ttgtgtgagt ctgttgtttg tttttttgttt 7740
ctttcccttt agcctttgtt ttctgtgtga aaccaaattg tcacagacat gtgccgccag 7800
caccccgaaaa aaaaaaacac catacccctg tcgattgtgt cgcccgtccc acgcggcgta 7860
tgcttaacct ttgagtgctg tacagtagcg agctaactct attaaattac atagcgtact 7920
ctaacggtag catgcgagtt tatctcatct actcatagtt gggtgtcctc ttcattccac 7980
agagtctgtt tccctacgct ttctacttct gtcgacgtgg caggggcga gacgatgacc 8040
gggatctgga gttcggcggt cccagccgac ggggaaccaa tggagcccaa cggaccggat 8100
ccggatccgc cgagatggcc aactgcgtca actccacggc ctcgtcggag atacacatga 8160
gcgtgatgcg tgcccgccag tatgccgtca atgtcacacc caccacggac gcccagatgc 8220
agcagctgca tcccctgtac accaactaaa cgaaaaacac atatggcgta tacttgattt 8280
caactccagt gcttcacaat gtctgttcat gagatgagtg cgcagtgcac agcaagtgtt 8340
ctccagaaaa ccagattgcc agcttacaaa tacaaatcac tcgataaaaa cattcctcaa 8400
aatggctgct aagagggcta aacaagaatt taaccacaaa taatagctaa agcaacagcc 8460
acagcaacaa tacgacaatc ccacagatta aataaagcct tcaagagcag caaattgacg 8520
tgcaaaataa attgaaatct tatagaagta tgtatttcga catacaatac atatttaggt 8580
gtcatggagt caggaatttc ataacgagat gtttataaat gcgcatgcaa atagatatat 8640
gaatttgaat gtttgatata cgtatataaa aagatagatg ttaacccaca aaaaaaaaaa 8700
taataaaata tgaaatatgc attgattggt aatctaaaac gaatcgcaaa ttaaagctac 8760
actgtgcaac agtggtatgt agtgagaaaa aaagtgtagg ttcaaggtct ggttgaaaga 8820
aaaaaatttt taatctgttc aattatttga aatattaact ttaaatacga tcaaacttat 8880
ctggtccttt cacattttat tcgattagcg aataaagact gacaatttta ttttatcagt 8940
gacttaaaaa tgatttaatt aaaacattat ccacatattc aaagtactta accaggcctt 9000
aaattcactt taactctcca ataaacaaag ataactaatg ttagttgtag cagagtgttc 9060
gtattgtttg taaaactcat actctcagtt ctgttagatg taatgacaaa gctttgattt 9120
gattcgttaa atacattaaa ataatctagt tcattgaaaa acaatcaaat tatatatgtc 9180
ataatactta catacatagg tattcatatt cataaagaat acatataggg cctaaagcgc 9240

acgactaac 9249

<210> 68
<211> 1675
<212> DNA
<213> Drosophila

<400> 68
ggtttggcaa acgattggca ataataggag tcatcaaaac gagagcaaat gtcatacagc 60
atgtcgtcct tattccgtgg cgttagtgcg acatttggag ttggaaacta ggagcccact 120
tagactaaca acgcccgtga ccattaccat gacgctactg cagaggcttc aggccatgtc 180
ggccaccacg accaggacaa tactggaggg cagcatcagc agttttggtg gcgggacaaa 240
tgagcctctg gcgagcaaaa tacccgttct ggaggagtca gcctcacatg ccagatattt 300
gaaattcatt gccgacgggc tcatcgacga gggactgggc agtgcggtgg gcagtgggag 360
cagcatcgcc gtatccgttg aagacgtggt cgccggacag gcgcaggaca tccaggcgag 420
cgaaggatcc accgacgacg ccgacggcag tagccatttg gcattagtct tcgtcaagtg 480
tttcattatt ggtttcatca tactggccgc catcctgggc aacatgctgg tgattgtgtc 540
ggtcatgcgg caccggaaat tgcgcatcat taccaactac tttgtggtct ctctggccgt 600
cgccgacatg ctggtggccc tctgtgcgat gacatttaat gcttccgtca tgatctcggg 660
caagtggatg tttggttccg tgatgtgcga catgtggaac agcttcgacg tctacttctc 720
caccgccagc atcatgcacc tctgttgcat atcggtcgac agatactacg ccattgtgca 780
gccactggac tatccactaa tcatgacaca gcgacgcgtg ttcatcatgc tattgatggt 840
gtggctatcg ccggcgctcc tctcgttcct gcccatctgc tcgggatggt acacaacgac 900
cgagaactac aagtatctca aatcgaatcc gcatatatgc gagttcaaag tgaacaaggc 960
atacgccata gtcagctcgt cgatgagctt ctggattccc ggcatcgtaa tgctgtcgat 1020
gtactaccgc atttaccagg aggccgaccg acaggagcgt ctggtgtaca gatccaaggt 1080
ggccgctctg ctgctggaga agcatctgca aattagccaa attcccaagc ccggccgag 1140
cattcaggtg gagcagtcga ccatctcgac gatgcggcgt gagcggaagg ccgcccgcac 1200
cctgggcatc atcatgagcg ccttcctcat ctgctggctg ccgttcttcc tctggtacat 1260
cgtatcctcg ctgtgcgata gttgcatcac tccgcgcctg ctcgttggca tcctgttttg 1320
gatcggctac ttcaactcgg ccctgaaccc cattatttat gcatacttca accgcgactt 1380
cagggccgcc ttcaagaaga ccctcaagag tctgtttccc tacgctttct acttctgtcg 1440
acgtggcagg gggcgagacg atgaccggga tctggagttc ggcggtccca gccgacgggg 1500
aaccaatgga gcccaacgga ccggatccgg atccgccgag atggccaact gcgtcaactc 1560
cacggcctcg tcggagatac acatgagcgt gatgcgtgcc cgccagtatg ccgtcaatgt 1620
cacacccacc acggacgccc agatgcagca gctgcatccc ctgtacacca actaa 1675

<210> 69
<211> 508
<212> PRT
<213> Drosophila

<400> 69
Met Thr Leu Leu Gln Arg Leu Gln Ala Met Ser Ala Thr Thr Thr Arg
1 5 10 15
Thr Ile Leu Glu Gly Ser Ile Ser Ser Phe Gly Gly Gly Thr Asn Glu
20 25 30
Pro Leu Ala Ser Lys Ile Pro Val Leu Glu Glu Ser Ala Ser His Ala
35 40 45
Arg Tyr Leu Lys Phe Ile Ala Asp Gly Leu Ile Asp Glu Gly Leu Gly
50 55 60
Ser Ala Val Gly Ser Gly Ser Ser Ile Ala Val Ser Val Glu Asp Val
65 70 75 80
Val Ala Gly Gln Ala Gln Asp Ile Gln Ala Ser Glu Gly Ser Thr Asp
85 90 95
Asp Ala Asp Gly Ser Ser His Leu Ala Leu Val Phe Val Lys Cys Phe
100 105 110
Ile Ile Gly Phe Ile Ile Leu Ala Ala Ile Leu Gly Asn Met Leu Val

```
        115                 120                 125
Ile Val Ser Val Met Arg His Arg Lys Leu Arg Ile Ile Thr Asn Tyr
    130                 135                 140
Phe Val Val Ser Leu Ala Val Ala Asp Met Leu Val Ala Leu Cys Ala
145                 150                 155                 160
Met Thr Phe Asn Ala Ser Val Met Ile Ser Gly Lys Trp Met Phe Gly
                165                 170                 175
Ser Val Met Cys Asp Met Trp Asn Ser Phe Asp Val Tyr Phe Ser Thr
            180                 185                 190
Ala Ser Ile Met His Leu Cys Cys Ile Ser Val Asp Arg Tyr Tyr Ala
        195                 200                 205
Ile Val Gln Pro Leu Asp Tyr Pro Leu Ile Met Thr Gln Arg Arg Val
    210                 215                 220
Phe Ile Met Leu Leu Met Val Trp Leu Ser Pro Ala Leu Leu Ser Phe
225                 230                 235                 240
Leu Pro Ile Cys Ser Gly Trp Tyr Thr Thr Thr Glu Asn Tyr Lys Tyr
                245                 250                 255
Leu Lys Ser Asn Pro His Ile Cys Glu Phe Lys Val Asn Lys Ala Tyr
            260                 265                 270
Ala Ile Val Ser Ser Ser Met Ser Phe Trp Ile Pro Gly Ile Val Met
        275                 280                 285
Leu Ser Met Tyr Tyr Arg Ile Tyr Gln Glu Ala Asp Arg Gln Glu Arg
    290                 295                 300
Leu Val Tyr Arg Ser Lys Val Ala Ala Leu Leu Leu Glu Lys His Leu
305                 310                 315                 320
Gln Ile Ser Gln Ile Pro Lys Pro Arg Pro Ser Ile Gln Val Glu Gln
                325                 330                 335
Ser Thr Ile Ser Thr Met Arg Arg Glu Arg Lys Ala Ala Arg Thr Leu
            340                 345                 350
Gly Ile Ile Met Ser Ala Phe Leu Ile Cys Trp Leu Pro Phe Phe Leu
        355                 360                 365
Trp Tyr Ile Val Ser Ser Leu Cys Asp Ser Cys Ile Thr Pro Arg Leu
    370                 375                 380
Leu Val Gly Ile Leu Phe Trp Ile Gly Tyr Phe Asn Ser Ala Leu Asn
385                 390                 395                 400
Pro Ile Ile Tyr Ala Tyr Phe Asn Arg Asp Phe Arg Ala Ala Phe Lys
                405                 410                 415
Lys Thr Leu Lys Ser Leu Phe Pro Tyr Ala Phe Tyr Phe Cys Arg Arg
            420                 425                 430
Gly Arg Gly Arg Asp Asp Asp Arg Asp Leu Glu Phe Gly Gly Pro Ser
        435                 440                 445
Arg Arg Gly Thr Asn Gly Ala Gln Arg Thr Gly Ser Gly Ser Ala Glu
    450                 455                 460
Met Ala Asn Cys Val Asn Ser Thr Ala Ser Ser Glu Ile His Met Ser
465                 470                 475                 480
Val Met Arg Ala Arg Gln Tyr Ala Val Asn Val Thr Pro Thr Thr Asp
                485                 490                 495
Ala Gln Met Gln Gln Leu His Pro Leu Tyr Thr Asn
            500                 505
```

<210> 70
<211> 3651
<212> DNA
<213> Drosophila

<400> 70
```
ctggacgccg gacacgaaag cagatcacct gccgtacagt gccaagatca tcgaacgcat 60
ttacaatgag gagatcggtg atggcggtgg ccacagtgca cggaggataa acatgctgga 120
gttcagtcag tatctggagc agtacctgtg gccacactac caacgtgaaa ccgccaccca 180
```

```
tgcccacctt atgtccatcg ttataatggc caacgagaag ttccgggagc gcgtcgaagt 240
ttggaccgtg tttgaaaagc tcccggatca gtatccagcc ttctttcgcc acgtgctaga 300
gagctgtttg ccgagcaaaa aggcaaagga ggctagtagc acgctaaggg agcggacagc 360
gctgttgatg ttcattaacc actgctttaa cagcatggag atagagctgt gcagagaaca 420
ggccaaacga ctagtctcct tgtccatgtg gcattgcctg cagcctcgta agttcttctt 480
cgcaactctt cacatctatc aaagtgatca aaaatgtatt tgcttgcagg ccgtcgcgaa 540
caagagcttc gagaagttcc tgagtggcga aagtactgga aacgcttgct caagaaggag 600
aaagacagca aaccggaggt cctctgggag agacacttta tgcagaacct tatcattgac 660
ttcctgcaca tactcgaaag cattcccgcc gaaggcgagg tgccgcgcaa cgtggttcac 720
tactgtgagc gcttcctaga gtttattatc gacttggagg cacttttgcc gactcgacgc 780
ttcttcaaca cggttttaga cgactgtcat ctgattgtgc gggctcttct gtcacccttg 840
gttcgtcgcg aggagggaaa actgtttggt caggtaggtg tcagacttct catttaataa 900
ctatctcgta cagcataaat tagttaaaat taagtacaaa tctcaatgtt ggacattgct 960
aaactacgaa aacacttata ttagtgtgga cggtggctcc tcagtaatcg ttgccgttca 1020
tataatgcag ctcatcgccc cagtcgtttg ccgacggtgg ctcattgtag cagcaggtct 1080
ttttcagcag aaatgtcaca tgtcgctggc gcagcacgca tatgtatagc aacagcggcg 1140
tctgcagggc gttaactacg atgtgggcgt aaagcaggcc ttccatctgc aaccaggaca 1200
ttatgaggaa tagccaggct atagacatta ctaacagcat cagcgagaac atgataaaac 1260
tgcaaggaag ataatgtttt catatactta tatgtaatgc ctagatgccg atctacttac 1320
ttggctttca atttgtgcgc aatgcggccg tataccgtcc tgcggttaat caacttcctg 1380
gtggtcacat agaagaatat gtttaccaaa atggtacaag caattggagc aaagaatatg 1440
catatgccta gccagccgat cgtctcttgc tcgcccacca tgtgctcctg tttgtaggat 1500
tcggcgtcga ggaagaaatg tgcgaagaca gccagtgcag ccatcgttgc tgtgcatccc 1560
caagcatacg cagaataata gcagtacttc ctgccgtccg tgactcgcag gaagacgttc 1620
ctcgaacgaa aggtcttcca aatgtaaaaa ccgaaactgt tcagccagaa gaaagctgcc 1680
aagaggctga agcacaggat aatgtccgcc acaatgaagc tcacatggct ggtcagctcg 1740
gtgaaaattc tgacaaggtc tgctgcctgg cttaccataa ggcacatggc tatcgttgtc 1800
acaatgttac caaccagatc tctgctgaat ggattgaaaa ttaacgcatg aagcaagtca 1860
aattcaaact attgctcacc tgagagtagg aagtataaaa tagatgatgg cgatgaccag 1920
caggataaca agcgagatgc catggaaaat tgggtttaag attttgcgga gcagaaagtt 1980
tgagtcactc cacttgatct ccttgcgggc caagcaaatg ttggcgaata gcacattttc 2040
ctctccagta ctggaagtgg cctgcgaatg gattttatga tcatccatgc aatttatgtg 2100
gactttagct actcttacct tgtcaatgca atacaatccc ttatcgtagt catggtatag 2160
cggttccaga ctgccggcaa tatcctcctc ataatccgat tgtattccat gacgctcctc 2220
tgcttcgttt tcggcattgg tgtaatgcct tagcctccca tcatccagta acactagctt 2280
atcagagctc tagaaaatat aaaaatatct ataagtatct aaaagtttct atgcaaagaa 2340
cttcaaactc acgccagcat agtggtatat tggccacatc tgcatggatc cacaatttgg 2400
aatcccgatg acgaatttga acttaacggg cctgttttgc tctcctccat agctggtgaa 2460
catcggctgg aagtagtctg caaatggaaa cggacgataa atagagaaca cagaccgtgg 2520
aaaatcataa aggcagttca agttaacgct gcttttgttg cggctggtgg atcagagggg 2580
agccagattt tcttttctct ctgcgaatat aaatatttac gcttctgggc aaagaacttg 2640
agctgttcca tgcgccttca tcagctggtt ttgcaatttg tgcagattgg ttcttcttca 2700
attatcggag ttcggataca ttttttttcca ctcagccatt catatgttaa tcatagctaa 2760
gcacacggaa attagattgc atgttgcgat tatgtgccct gttgtattgt agtattggga 2820
ttactcaaat gactacaatt aaaatcggat gagatcagta caatgcaatg attcaactat 2880
cagtaatagt tacattttta tgtgtcacct atatggccga tccatggtca taaatgtttt 2940
atccagtttt agttgtccga gctagaatat cgatatagat gctagagtac gcttttctat 3000
ttcccacaat tgttgcgatt gtttgtcaaa cattcaccca aggcaatata tcactatatt 3060
caccattagc tccacaaaca ttaattacca gggctcgcat ttatggggaa cagaagcatc 3120
ttatattacc ttactaatgg attttttgcc tcattttact tccaagagat aatattcccg 3180
atttgggaat ccctcatcaa gtggctactt tatagtttaa acagaaacat gccccacagc 3240
ataaatcata tttctcatca ggattggggc caccctctgt ttgttttaag aactttcttg 3300
ggggagtgcg aactgtggag ttggggatat ttagctgttg gcttcgttcc actgatccac 3360
tgatagccaa caaatgggcc atttaatgtt gggcatacat atgttgcatg ccagaaggga 3420
aaatccataa agttgaatgg ccaaccaact tcgttgtggt tgttatggca actttttcga 3480
gcgactagga gttttccgcc tggccagctg ttgaacttac ccgtttcatt gacctgctga 3540
cactcatggt ccacgtgtat ctcgaacttc tcgcaacatt tgttgaccag cactaggttg 3600
ggatcggaag atagggctgt tgatgatccg gcgctagtga catgactggt t          3651
```

<210> 71

<211> 1272
<212> DNA
<213> Drosophila

<400> 71
atggaacagc tcaagttctt tgcccagaag cactacttcc agccgatgtt caccagctat 60
ggaggagagc aaaacaggcc cgttaagttc aaattcgtca tcgggattcc aaattgtgga 120
tccatgcaga tgtggccaat ataccactat gctggcagct ctgataagct agtgttactg 180
gatgatggga ggctaaggca ttacaccaat gccgaaaacg aagcagagga gcgtcatgga 240
atacaatcgg attatgagga ggatattgcc ggcagtctgg aaccgctata ccatgactac 300
gataagggat tgtattgcat tgacaaggcc acttccagta ctggagagga aaatgtgcta 360
ttcgccaaca tttgcttggc ccgcaaggag atcaagtgga gtgactcaaa ctttctgctc 420
cgcaaaatct taaacccaat tttccatggc atctcgcttg ttatcctgct ggtcatcgcc 480
atcatctatt ttatacttcc tactctcagc agagatctgg ttggtaacat tgtgacaacg 540
atagccatgt gccttatggt aagccaggca gcagaccttg tcagaatttt caccgagctg 600
accagccatg tgagcttcat tgtggcggac attatcctgt gcttcagcct cttggcagct 660
ttcttctggc tgaacagttt cggtttttac atttggaaga cctttcgttc gaggaacgtc 720
ttcctgcgag tcacggacgg caggaagtac tgctattatt ctgcgtatgc ttggggatgc 780
acagcaacga tggctgcact ggctgtcttc gcacatttct tcctcgacgc cgaatcctac 840
aaacaggagc acatggtggg cgagcaagag acgatcggct ggctaggcat atgcatattc 900
tttgctccaa ttgcttgtac cattttggta aacatattct tctatgtgac caccaggaag 960
ttgattaacc gcaggacggt atacggccgc attgcgcaca aattgaaagc caattttatc 1020
atgttctcgc tgatgctgtt agtaatgtct atagcctggc tattcctcat aatgtcctgg 1080
ttgcagatgg aaggcctgct ttacgcccac atcgtagtta acgccctgca gacgccgctg 1140
ttgctataca tatgcgtgct gcgccagcga catgtgacat ttctgctgaa aaagacctgc 1200
tgctacaatg agccaccgtc ggcaaacgac tggggcgatg agctgcatta tatgaacggc 1260
aacgattact ga 1272

<210> 72
<211> 423
<212> PRT
<213> Drosophila

<400> 72
Met Glu Gln Leu Lys Phe Phe Ala Gln Lys His Tyr Phe Gln Pro Met
1 5 10 15
Phe Thr Ser Tyr Gly Gly Glu Gln Asn Arg Pro Val Lys Phe Lys Phe
20 25 30
Val Ile Gly Ile Pro Asn Cys Gly Ser Met Gln Met Trp Pro Ile Tyr
35 40 45
His Tyr Ala Gly Ser Ser Asp Lys Leu Val Leu Leu Asp Asp Gly Arg
50 55 60
Leu Arg His Tyr Thr Asn Ala Glu Asn Glu Ala Glu Glu Arg His Gly
65 70 75 80
Ile Gln Ser Asp Tyr Glu Glu Asp Ile Ala Gly Ser Leu Glu Pro Leu
85 90 95
Tyr His Asp Tyr Asp Lys Gly Leu Tyr Cys Ile Asp Lys Ala Thr Ser
100 105 110
Ser Thr Gly Glu Glu Asn Val Leu Phe Ala Asn Ile Cys Leu Ala Arg
115 120 125
Lys Glu Ile Lys Trp Ser Asp Ser Asn Phe Leu Leu Arg Lys Ile Leu
130 135 140
Asn Pro Ile Phe His Gly Ile Ser Leu Val Ile Leu Leu Val Ile Ala
145 150 155 160
Ile Ile Tyr Phe Ile Leu Pro Thr Leu Ser Arg Asp Leu Val Gly Asn
165 170 175
Ile Val Thr Thr Ile Ala Met Cys Leu Met Val Ser Gln Ala Ala Asp
180 185 190
Leu Val Arg Ile Phe Thr Glu Leu Thr Ser His Val Ser Phe Ile Val
195 200 205

```
Ala Asp Ile Ile Leu Cys Phe Ser Leu Leu Ala Ala Phe Phe Trp Leu
    210                 215                 220
Asn Ser Phe Gly Phe Tyr Ile Trp Lys Thr Phe Arg Ser Arg Asn Val
225                 230                 235                 240
Phe Leu Arg Val Thr Asp Gly Arg Lys Tyr Cys Tyr Tyr Ser Ala Tyr
                245                 250                 255
Ala Trp Gly Cys Thr Ala Thr Met Ala Ala Leu Ala Val Phe Ala His
            260                 265                 270
Phe Phe Leu Asp Ala Glu Ser Tyr Lys Gln Glu His Met Val Gly Glu
        275                 280                 285
Gln Glu Thr Ile Gly Trp Leu Gly Ile Cys Ile Phe Phe Ala Pro Ile
    290                 295                 300
Ala Cys Thr Ile Leu Val Asn Ile Phe Phe Tyr Val Thr Thr Arg Lys
305                 310                 315                 320
Leu Ile Asn Arg Arg Thr Val Tyr Gly Arg Ile Ala His Lys Leu Lys
                325                 330                 335
Ala Asn Phe Ile Met Phe Ser Leu Met Leu Leu Val Met Ser Ile Ala
            340                 345                 350
Trp Leu Phe Leu Ile Met Ser Trp Leu Gln Met Glu Gly Leu Leu Tyr
        355                 360                 365
Ala His Ile Val Val Asn Ala Leu Gln Thr Pro Leu Leu Leu Tyr Ile
    370                 375                 380
Cys Val Leu Arg Gln Arg His Val Thr Phe Leu Leu Lys Lys Thr Cys
385                 390                 395                 400
Cys Tyr Asn Glu Pro Pro Ser Ala Asn Asp Trp Gly Asp Glu Leu His
                405                 410                 415
Tyr Met Asn Gly Asn Asp Tyr
            420


<210> 73
<211> 5299
<212> DNA
<213> Drosophila

<400> 73
gctgttcgca cggcattttt gggccagcaa acaggatgcg gaaaaccgga agctgggaac 60
tggaaatggg caagaatctg gtttttggct gccattcaat taggtatcgg tcagcccaat 120
tcaactgctg ctaactggca actggagatt tggcgattgc agctgcccat tgtcaacact 180
tcaattaacg caatagcagg taaaaagggg acatcttttc atcgagatct cgaccttgtt 240
ttgctatcct tgcaaacaat tgatggtatg tcgtaagttt gtgcgaataa aatggttttg 300
gaaatatata tatatattgt ttattagttt atggccccaa aaaatgctac aacacaatat 360
tgaaaccgtt gttatctgaa gggaatgcaa cgattcgagc ttcatattga gaaggtcctg 420
ttttctctct gttcctgcat agttcctctt tcatttcgca tctttttcca ctggccgcag 480
tctcgcagga tatgctaact aaggactatg ccaccatatt cgacaatttg ttgggcacca 540
agttaatatg cattgatgcg ttattcatac gccccgtggg acgggcgaat agcaaaactc 600
tcaatgggcc tttgttaatt tgtatgtgtg tgtatgtgta ttttgcgtgt atttgatttg 660
atgcgtttag tacttaggat aacaattacg tttaggttaa ctagcataaa tctataatca 720
tcacattatg gcacaatgta tgttatgtag tatatatata tcgtatatat agagagactg 780
atatacatag atatcgattg atataacaaa ttaaattttc agctaaccgc acgcttcaac 840
tgaatcactt ttgattcaat catctcactt gtgattccgc gtgatttgct ggcaaattta 900
aagccctctt ctcgtgtttc atgagaagtt tcatgatatc atctatcact cctctcatgg 960
agtttaaaac cataaatatt ttattataaa tattcactat tcacgctatc aggcgacccg 1020
tattacgatt tcgacgactt ccggctccgc cataattact accataatga tcgccattta 1080
tttgggtcag tgtagtcgtt gtatgcggca catagagact cgtattgtta cgcgcatgtc 1140
gacgtctata gcgatatccc gttatcagtt ctcgaaccgt ctttcggtac ttatcggaaa 1200
gtacacagta tagaaaaaaa ttggatgccg catgtaggga ggccaatagg ttacacacat 1260
ttcccgccac cacctgagtg ccaatttgaa tttccatata ctgattgact atcaaataga 1320
tcgcccaagg caattggcaa actatgaaca ttaacaccac cgcaataagt gtgatcgtca 1380
cgcgattttc ctgcgatacg ctgcctttta gtccagaatt tgatttgcgc taaggattca 1440
```

gaatttatac aacatatcat catcatcatc atcgtttcga ataactcact tttgtgcgcc 1500
ttatgctgct ggcattggtc aggtcaccac gcagattctt ggaccgatga accaatagaa 1560
tgagaatgga gttgaaggtg gccaatagaa gtagggggaa taccacgaag accagcaggg 1620
tgaacaggtt ccacagatag ctgtatagct catcatagcc cagttccgaa agaccatgat 1680
tgtaataggt cacctaaaac atatagatag atacatagaa tacatacata gatttccatt 1740
aagaacaacc ttttaagcaa tggaatatcc caagcgacta agtaaaagca tgtcttacaa 1800
aacccaaagc actgatatca ttttaagtta tcaccgaatt tatcatagtc atggcgcttg 1860
tacaccgctc gtgattttca ctggttaaca attttctcat gtttttttct ccagggaaat 1920
cgacggtatt tccaacatta ttttactcac attttgacag tattccgtta cgttaaatgc 1980
gaatgcatcc gtattgttga tgttgtgatt ctcagcgctg cgcttgcgac gccacaattg 2040
cagcaaactc tctgtgaccc ctgacctctg ggccgcgtct gccaccttct gatcctgctc 2100
ctccagcgtt tttctcagtg tgggcagcgt cacaaatcca gatgactgcc agtgacggcg 2160
atgtcgtgga atgttctcta gttgagggca aagaaagtaa tgcaactatc aactacttca 2220
cttataaagc tgtctaaaag tatgcttcgt acttttcaat gcatacttct aagcatctct 2280
attaatacaa cttcgatttt tccaagtgta cactactgct ataccctcgt tactcacctg 2340
gctctccatc accactgcca ctgccccagt cgacgagctg attagagctg ccggacgtcg 2400
tggactccat ggcggagtcc ggcgattgcc aaatggtcgc cggcgtgggc agcgggggcg 2460
tggccagtgg gggcgtgact gccacgggtg gcggcggcat gggcgagatg ttggccaccg 2520
tttggttgca cggctggtag gcgtcatcaa tctgcctagt accaatcgta attgtatgct 2580
cgaatgccgt cgatagtgtg gacagcaaac agaagatggc caccgctggc aatgtagaat 2640
ataaagttgc atatggaaat ggattctatt gttacatacg cctatttcag ccaactcacc 2700
tgctatcacc tttttggcca gcgattccgt gcagaatgtt tgcctcttga gcggatagcg 2760
tatcgcaatg aatctctcga tggtgaagca cacggctatg taaattgaga tgtaacctgt 2820
ccggtaaaaa cagagaaagc aattgtgaat gggggcgtgg cagtacaata cgaacgcaat 2880
caaagatgca ctgcaaatat tttgggaaaa tccaatccaa aaaggatata tatacatata 2940
tatttattca ggttaagtgt ctaatataga ttggtaaatt aaattcaaac ttcggttaga 3000
ctgtgccttt tttccgtgta ctctggaatt ataaataaac gcagctccgc tgttaatacc 3060
ccaaaccatc cattttttgac cgcaccacca atttatttgc cgcgcgggct gtcaatttat 3120
tcaaaacgca gcgcactcta agcgattgac gtccgtgctg tgctgatgag gttgctgcgg 3180
gagggtggga tattcaaggg gggtgagggg ggtgaaaagt ggggtgggaa agcggaaaac 3240
ggttcggaaa aacgctatcg catggcagtg agcctcagct gaccgctggc atttcgcaga 3300
aagcgctgct tactatacct cgtaaattgg cagagctcaa tgaaattgat gcacaacaga 3360
aacagcagca gcagcagcag cagcagcagc agcagcagca acagccgcag gtgcaattcg 3420
aattgatggg cccatcgctg gaatccggtt gatacggggg gaaatggggg atacattgga 3480
tccggggatc cagacgaggc taacccgggg aactggggaa tttatggaat ccgcaccatt 3540
ccacttagtg ctctccaaac gttgcattcg aactgttttt gtttttttgcc cccggcagca 3600
attgttttca ccgcatgata tgtacgcttt tcttattgta gttttgcggg aataatgtta 3660
aatataaaaa caccctttttg gtaaattttt tcaaattgaa agcttattta agcataaaat 3720
actcgtgttt cgatttaaag aacactaatc gagttacttt cggtctcaca accttattta 3780
ttgcatactt tgcgactgct ttgaattatt gaattattac tccattgcta cttaattcta 3840
atttggtcag tactcaccga aactgtcgca cagccagaca aaatagccat agagctgcca 3900
gtagagcttg aaatgatact tgggatagtc atagtgctgg agtgacagga ttaactggca 3960
cgtcaggtag gcaatatccg taatggccag tgccgttaga tatatgttcg tcgtgcatct 4020
catgcgcttt ctaaaaatac cgtcaagaat atcgaaaaga attgcaaatg gtgttgattc 4080
tgttgggcac tcgacgcatt tggaaatcag gccaaaccat ttcggtggac aatgcaaaaa 4140
cgctgccaag tgcaacacca attgaaatag caaattgcct gaaagccaaa cgatcggtgg 4200
ggttgtgcgc agagggggtg gaaaattcca ggcggttgat gtgatgggcg cggggccttg 4260
ggcattttcc gaaatgcgca caaatgtcgg cacagtcata aagccaaaag atgaacaata 4320
aaaactaact gttagcattt agttttctcc ctggcatcgc ccgcagctct gatagttaaa 4380
ttctcttgta ttcccgtaaa acaaaataaa ataacgtaat gaaaatattt aattaaaaag 4440
ccattcatgc cgtcaacgct caatgttaag ctagttgctt tgaaaagttg taaatatggt 4500
ttaaaataat aaagttcatg ctataaatat caataaattt tgtttagcaa ataaatgaat 4560
ttgctaagat cttagcaagg aacacagttt atcttttttt ttcacgtttt cctgggaaat 4620
ggacatatgg tgtcatagca gtcagcctaa cgcacagagc acttgcaact gtcgcatatt 4680
ttccggcaga cataaaaaac acacccatcc acacagcata tcgcctatct gtgtgcgttt 4740
ttgcatttgg cacataaaat ggctcttatt gtttcagtta ctgttattgt tgttgcactt 4800
gtacccaacg tagttgttgc tgttgttatt gttatcgttg ttggctgtgt ggaaatgatt 4860
ttaaagtgca gcggtggcta caattaaatg tcgtgtttaa tgcgctgcgt acctgctcgc 4920
cgcttcatcc atgttccact tatttcacac acacagacac acacacacac gcacaacagg 4980
acatgcagtt ggatgcacag agagaaaatc gtgcaagtta gctggcattt taaatggaaa 5040

attgtaactt aaaactagtt taaaacaagc tcaaacgaat aagttgaagt ttggaataaa 5100
attaatttga aaaggctcat atattttata tgcattgcat actttttagt acttaaccaa 5160
atgcatgaaa atccaattag tcgcttttgc atacctaata acttaacatt cttaacgtgg 5220
cagctaaaac ttctttcaaa tgccatcatc agctaaatga ttcatattta attcgaaaat 5280
cttgtgaccc tgctgcatg 5299

<210> 74
<211> 1467
<212> DNA
<213> Drosophila

<400> 74
atgactgtgc cgacatttgt gcgcatttcg gaaaatgccc aaggccccgc gcccatcaca 60
tcaaccgcct ggaattttcc accccctctg cgcacaaccc caccgatcgt ttggctttca 120
ggcaatttgc tatttcaatt ggtgttgcac ttggcagcgt tttgcattg tccaccgaaa 180
tggtttggcc tgatttccaa atgcgtcgag tgcccaacag aatcaacacc atttgcaatt 240
cttttcgata ttcttgacgg tatttttaga aagcgcatga gatgcacgac gaacatatat 300
ctaacggcac tggccattac ggatattgcc tacctgacgt gccagttaat cctgtcactc 360
cagcactatg actatcccaa gtatcatttc aagctctact ggcagctcta tggctatttt 420
gtctggctgt gcgacagttt cggttacatc tcaatttaca tagccgtgtg cttcaccatc 480
gagagattca ttgcgatacg ctatccgctc aagaggcaaa cattctgcac ggaatcgctg 540
gccaaaaagg tgatagcagc ggtggccatc ttctgtttgc tgtccacact atcgacggca 600
ttcgagcata caattacgat tggtactagg cagattgatg acgcctacca gccgtgcaac 660
caaacggtgg ccaacatctc gcccatgccg ccgccacccg tggcagtcac gccccccactg 720
gccacgcccc cgctgcccac gccggcgacc atttggcaat cgccggactc cgccatggag 780
tccacgacgt ccggcagctc taatcagctc gtcgactggg gcagtggcag tggtgatgga 840
gagccagaga acattccacg acatcgccgt cactggcagt catctggatt tgtgacgctg 900
cccacactga gaaaaacgct ggaggagcag gatcagaagg tggcagacgc ggcccagagg 960
tcaggggtca cagagagttt gctgcaattg tggcgtcgca agcgcagcgc tgagaatcac 1020
aacatcaaca atacggatgc attcgcattt aacgtaacgg aatactgtca aaatgtgacc 1080
tattacaatc atggtctttc ggaactgggc tatgatgagc tatacagcta tctgtggaac 1140
ctgttcaccc tgctggtctt cgtggtattc cccctacttc tattggccac cttcaactcc 1200
attctcattc tattggttca tcggtccaag aatctgcgtg gtgacctgac caatgccagc 1260
agcataaggc gcacaaaata ccgaaagacg gttcgagaac tgataacggg atatcgctat 1320
agacgtcgac atgcgcgtaa caatacgagt ctctatgtgc cgcatacaac gactacactg 1380
acccaaataa atggcgatca ttatggtagt aattatggcg gagccggaag tcgtcgaaat 1440
cgtaatacgg gtcgcctgat agcgtga 1467

<210> 75
<211> 488
<212> PRT
<213> Drosophila

<400> 75
Met Thr Val Pro Thr Phe Val Arg Ile Ser Glu Asn Ala Gln Gly Pro
1 5 10 15
Ala Pro Ile Thr Ser Thr Ala Trp Asn Phe Pro Pro Pro Leu Arg Thr
20 25 30
Thr Pro Pro Ile Val Trp Leu Ser Gly Asn Leu Leu Phe Gln Leu Val
35 40 45
Leu His Leu Ala Ala Phe Leu His Cys Pro Pro Lys Trp Phe Gly Leu
50 55 60
Ile Ser Lys Cys Val Glu Cys Pro Thr Glu Ser Thr Pro Phe Ala Ile
65 70 75 80
Leu Phe Asp Ile Leu Asp Gly Ile Phe Arg Lys Arg Met Arg Cys Thr
85 90 95
Thr Asn Ile Tyr Leu Thr Ala Leu Ala Ile Thr Asp Ile Ala Tyr Leu
100 105 110

Thr Cys Gln Leu Ile Leu Ser Leu Gln His Tyr Asp Tyr Pro Lys Tyr
115 120 125
His Phe Lys Leu Tyr Trp Gln Leu Tyr Gly Tyr Phe Val Trp Leu Cys
130 135 140
Asp Ser Phe Gly Tyr Ile Ser Ile Tyr Ile Ala Val Cys Phe Thr Ile
145 150 155 160
Glu Arg Phe Ile Ala Ile Arg Tyr Pro Leu Lys Arg Gln Thr Phe Cys
165 170 175
Thr Glu Ser Leu Ala Lys Lys Val Ile Ala Ala Val Ala Ile Phe Cys
180 185 190
Leu Leu Ser Thr Leu Ser Thr Ala Phe Glu His Thr Ile Thr Ile Gly
195 200 205
Thr Arg Gln Ile Asp Asp Ala Tyr Gln Pro Cys Asn Gln Thr Val Ala
210 215 220
Asn Ile Ser Pro Met Pro Pro Pro Pro Val Ala Val Thr Pro Pro Leu
225 230 235 240
Ala Thr Pro Pro Leu Pro Thr Pro Ala Thr Ile Trp Gln Ser Pro Asp
245 250 255
Ser Ala Met Glu Ser Thr Thr Ser Gly Ser Ser Asn Gln Leu Val Asp
260 265 270
Trp Gly Ser Gly Ser Gly Asp Gly Glu Pro Glu Asn Ile Pro Arg His
275 280 285
Arg Arg His Trp Gln Ser Ser Gly Phe Val Thr Leu Pro Thr Leu Arg
290 295 300
Lys Thr Leu Glu Glu Gln Asp Gln Lys Val Ala Asp Ala Ala Gln Arg
305 310 315 320
Ser Gly Val Thr Glu Ser Leu Leu Gln Leu Trp Arg Arg Lys Arg Ser
325 330 335
Ala Glu Asn His Asn Ile Asn Asn Thr Asp Ala Phe Ala Phe Asn Val
340 345 350
Thr Glu Tyr Cys Gln Asn Val Thr Tyr Tyr Asn His Gly Leu Ser Glu
355 360 365
Leu Gly Tyr Asp Glu Leu Tyr Ser Tyr Leu Trp Asn Leu Phe Thr Leu
370 375 380
Leu Val Phe Val Val Phe Pro Leu Leu Leu Leu Ala Thr Phe Asn Ser
385 390 395 400
Ile Leu Ile Leu Leu Val His Arg Ser Lys Asn Leu Arg Gly Asp Leu
405 410 415
Thr Asn Ala Ser Ser Ile Arg Arg Thr Lys Tyr Arg Lys Thr Val Arg
420 425 430
Glu Leu Ile Thr Gly Tyr Arg Tyr Arg Arg Arg His Ala Arg Asn Asn
435 440 445
Thr Ser Leu Tyr Val Pro His Thr Thr Thr Thr Leu Thr Gln Ile Asn
450 455 460
Gly Asp His Tyr Gly Ser Asn Tyr Gly Gly Ala Gly Ser Arg Arg Asn
465 470 475 480
Arg Asn Thr Gly Arg Leu Ile Ala
485

<210> 76
<211> 3946
<212> DNA
<213> Drosophila

<400> 76
ataaatggcc gaatttgcaa gtcagttcca cttgctcttt aaaaagcgaa tctcaccgaa 60
atttaagcaa aatttatcac cgaccaatgt gaaattccac taataatttt tcagcggtgt 120
tgatggcttt attgacatcg ccagaaatga ttcgagctgt agcagcgaaa atttaatcaa 180
agacctaacc cacgaacctg gacgttctca catagaaaat gaaaacgaaa gctgaaagac 240

agatggatcg ttcgatgtga gctcttttta cacttggctt acacgacgac tgcggcgtcg 300
gtaaaattaa tataccaaac aattacacgg cccggcaaaa caattttaag cccagttaat 360
gccgtcgcac agtggctgca agccacgcaa attcgccgcg aaaaattgcc cacaattgcc 420
agtggccgag aaatgaaagc gctcaaattt gcatactcgt actgtcaaat gaagcgtagc 480
aaaaggaaaa ttgctacaca tactaattgg aaaagaaatg caaaaccaat gcaaatgaaa 540
ataatgacag ggcgcggttt gcgaagctgt ctaattattt atttgacatt ggcaatgaca 600
atgaaaccga gagtcccaca gcagagtaat ggatttttct ttcaactgct taagttcgtt 660
ttccattgaa atcttatttg gacatataaa ccatacgtag aatgataaga ataataaaaa 720
acaagtatgt tggacacaat tgataaggaa tctttagtac tttctcacct aaaaaaggtg 780
atatatattt gaattcatgt tgtcatgata gctactcgaa ccactgtgcg ttggctacag 840
tctgctgatg ctcacagtgt ttgaaagtta agacttttgc tttccttccc ccgacttggt 900
ttttgctttt ggaaagacca agcaaatgga taaacaagag ttcaagaggt tgtcagcccc 960
gcccctttgc gcacatgagg caagaaagaa aagaaagtta ttacaaaatt aacagaggct 1020
gacagctgac aagggcagag gaatggggcc gaaatggaac tgtatctagc ttgagttggg 1080
tccgaccgcc tcgggcagaa aatgtgacaa atttggaggg gaaaagctcg tatgagcggg 1140
acttaccacc tccattgttg gcactggagc agtgtttgga gctgagtccc tggcccgtgt 1200
tggatgcaga gttgggcagt ccgcagccgc tcatcacgcc gtacttgatg ggcttgaagt 1260
ggaagaacga gggcgagtag ccggaaccgg cgcgactcag ggaactgaac tgctcctcgc 1320
ggtcctccac gtacaggccc tccttgccca tggccgccag ttgcctgccc ttgggcatga 1380
acatcaccaa gaagaccgtg gcggatgtgg ctaccagacc gaaggccacg caagcgtcct 1440
tgtgccgctc ggccacggcc aggccgcata gcatccatcc cagccagatg gggatggctc 1500
ctccgatggc cagcccaatg taggtggcct cacggtagtt gtccctgatg ccgcgcgact 1560
tgatggctag cacggcgatg aagacgatca gaaagacaat gtagatcagc gagaagagca 1620
gctccgagaa ctgcgtcttg cagagcggta ttagtacggt gctaaccgcc gctatccggg 1680
tgtagatctc cggagtcccg tccagcgttg tgtaggacgt ggggtaaaac agtgccgagt 1740
agttggtctg tgaggccacc gtggtgctaa ggaagccact gcccatcaca ggcacactgg 1800
tggtgtatac ctccggtggt tgggtgagca gccactgccc tccgatcgct acctggatga 1860
gcagcgcaaa gaggagcagc aatccctggt atggagccgg cagatacacc cctccattta 1920
gactaatcag aaatacgcat ttgaccagca gggcggcaaa gaccagggcg taggccacgc 1980
ccactccgaa gcggatggcc ccgcaactga tcaacgaggg ctgtgcggtg atgatggccc 2040
ccaaactagc gcaggcgaac agaccgagaa gcagcatctg gcccagaaag aggtgcctac 2100
gcgacgggga cgtccgccag gccttgaaga gcacaaagat ctcgaacgct gccatcatga 2160
gcatggtgag ggtggccagg actagcaccg gcaccaccca cggctccttg cgtagtcccg 2220
cgaagggcaa gtgggtggac actatactgg gcggtgtgct gttggagctg ctcccgggag 2280
tcacccgtcc agtggtgccg gctgctcgcg ccttcgcccc tgaactccga cggggcgtgg 2340
ctggtgtggt ggtggacgtg gtgctgctag aactgctggc cgatagtgtg ggtaccacaa 2400
agaattccgt cgacatttcg agacgggcag cactggagac attgccaccc ggaaaattga 2460
gatccaattg tttggcagat cgctgggtct gaggagctgc cggagtagtg gtactagtgg 2520
tgctacttgt agtacttgta gtcgtgctgg tagtagtggt ggtgctggtg gcattacggc 2580
tactcggctt aagcacctgg taggcgttgc gatccgattc ctggctggtc accaggcgca 2640
cccgccgcca aatgttgttg ccattaacgc tgaggctgcg cgacgccggc gtggtggcga 2700
tcaactgcat gatcagcagc aggatcagca gcaggtggcg tctaccggcc attggatgtg 2760
gtgaccttcg atgccctgga cctctgacaa tataatcaaa cgcggcacat ctcttgcaaa 2820
cgcacactaa ccacgcactc gcgatcgcga ttgttcaggt tcgtaaattc ttcagtcctc 2880
gcaggcccta gatcatgaat tcttcttctc gattctcgcg ctaaaacgtc cgctcgctcc 2940
gatcgaccaa ctgcaactga ctggcggctc tgcaagttgg gctttggacc ggcttagtcg 3000
tttgagacca gtttgacggc agccaagtcg aactaacact tgtttgagtt cgatgtgcgt 3060
ttgctaaccc ggctaacact tacaggaacg caacatgatg cgtgggaggc agttggggtt 3120
tggggcgctt tatatgattt tttgcaatgt agagtattgc acgagcattt actaaagcat 3180
cctttgaatt tgatacaaca ttcgttactc gatcttctgg ctttaaggtt tagtatatag 3240
cttcaacaac taatgttcta aatagctcaa tgtcttgctg aagggctatt gatttcctct 3300
cgaaagctat ggtacttaat ttcctattta gagccagtaa gaactcaatt tatctttagc 3360
ttccgtgcta atccgctgga acatatactt taacacaaca tttacatacg taagtttttt 3420
ttccatttga taattattcc aattagattt catattgcgg gatattaaca ttggacacaa 3480
tttttttcctc tgcggttggt tgcaacatca actgcagacg cgatccacag acctagctca 3540
ggatccacac tcaaaccgac tgcgggatgg agaacacttt cattcgttct gcatatgcgg 3600
aatttcggtt tatgggctct gatgctcgag ttggctggta ggaaagtgct atgggccatt 3660
atctattagc caaccgcttt tacggtctcc tcgcggagcc tctctcctcg aaaacgagcc 3720
aacaaaaggc caaatagcga aatccgagca caaaaggtga ttgtaaatca aaagagcgct 3780
aatggtatat cgatatacaa agtgttgagg cgggaaaggc atcgccaata gcgttagaaa 3840

cggcaatagc aatagcaacc cgctctgatt ccttaggtct tattgtattc ggcgcgcgct 3900
ttcagtcaac cgcagttaga attcagatca cctgaatgag tgggca 3946

<210> 77
<211> 1946
<212> DNA
<213> Drosophila

<400> 77
tcgatcggag cgagcggacg ttttagcgcg agaatcgaga agaagaattc atgatctagg 60
gcctgcgagg actgaagaat ttacgaacct gaacaatcgc gatcgcgagt gcgtggttag 120
tgtgcgtttg caagagatgt gccgcgtttg attatattgt cagaggtcca gggcatcgaa 180
ggtcaccaca tccaatggcc ggtagacgcc acctgctgct gatcctgctg ctgatcatgc 240
agttgatcgc caccacgccg gcgtcgcgca gcctcagcgt taatggcaac aacatttggc 300
ggcgggtgcg cctggtgacc agccaggaat cggatcgcaa cgcctaccag gtgcttaagc 360
cgagtagccg taatgccacc agcaccacca ctactaccag cacgactaca agtactacaa 420
gtagcaccac tagtaccact actccggcag ctcctcagac ccagcgatct gccaaacaat 480
tggatctcaa ttttccgggt ggcaatgtct ccagtgctgc ccgtctcgaa atgtcgacgg 540
aattctttgt ggtacccaca ctatcggcca gcagttctag cagcaccacg tccaccacca 600
caccagccac gccccgtcgg agttcagggg cgaaggcgcg agcagccggc accactggac 660
gggtgactcc cgggagcagc tccaacagca caccgcccag tatagtgtcc acccacttgc 720
ccttcgcggg actacgcaag gagccgtggg tggtgccggt gctagtcctg gccaccctca 780
ccatgctcat gatggcagcg ttcgagatct ttgtgctctt caaggcctgg cggacgtccc 840
cgtcgcgtag gcacctcttt ctgggccaga tgctgcttct cggtctgttc gcctgcgcta 900
gtttgggggc catcatcacc gcacagccct cgttgatcag ttgcggggcc atccgcttcg 960
gagtgggcgt ggcctacgcc ctggtctttg ccgccctgct ggtcaaatgc gtatttctga 1020
ttagtctaaa tggagggggtg tatctgccgg ctccatacca gggattgctg ctcctctttg 1080
cgctgctcat ccaggtagcg atcggagggc agtggctgct cacccaacca ccggaggtat 1140
acaccaccag tgtgcctgtg atgggcagtg gcttccttag caccacggtg gcctcacaga 1200
ccaactactc ggcactgttt taccccacgt cctacacaac gctggacggg actccggaga 1260
tctacacccg gatagcggcg gttagcaccg tactaatacc gctctgcaag acgcagttct 1320
cggagctgct cttctcgctg atctacattg tctttctgat cgtcttcatc gccgtgctag 1380
ccatcaagtc gcgcggcatc agggacaact accgtgaggc cacctacatt gggctggcca 1440
tcggaggagc catccccatc tggctgggat ggatgctatg cggcctggcc gtggccgagc 1500
ggcacaagga cgcttgcgtg gccttcggtc tggtagccac atccgccacg gtcttcttgg 1560
tgatgttcat gcccaagggc aggcaactgg cggccatggg caaggagggc ctgtacgtgg 1620
aggaccgcga ggagcagttc agttccctga gtcgcgccgg ttccggctac tcgccctcgt 1680
tcttccactt caagcccatc aagtacggcg tgatgagcgg ctgcggactg cccaactctg 1740
catccaacac gggccaggga ctcagctcca aacactgctc cagtgccaac aatggaggtg 1800
gtaagtcccg ctcatacgag cttttcccct ccaaatttgt cacattttct gcccgaggcg 1860
gtcggaccca actcaagcta gatacagttc catttcggcc ccattcctct gcccttgtca 1920
gctgtcagcc tctgttaatt ttgtaa 1946

<210> 78
<211> 583
<212> PRT
<213> Drosophila

<400> 78
Met Ala Gly Arg Arg His Leu Leu Leu Ile Leu Leu Leu Ile Met Gln
1 5 10 15
Leu Ile Ala Thr Thr Pro Ala Ser Arg Ser Leu Ser Val Asn Gly Asn
20 25 30
Asn Ile Trp Arg Arg Val Arg Leu Val Thr Ser Gln Glu Ser Asp Arg
35 40 45
Asn Ala Tyr Gln Val Leu Lys Pro Ser Ser Arg Asn Ala Thr Ser Thr
50 55 60
Thr Thr Thr Thr Ser Thr Thr Thr Ser Thr Thr Ser Ser Thr Thr Ser
65 70 75 80
Thr Thr Thr Pro Ala Ala Pro Gln Thr Gln Arg Ser Ala Lys Gln Leu

```
                    85                      90                      95
Asp Leu Asn Phe Pro Gly Gly Asn Val Ser Ser Ala Ala Arg Leu Glu
                100                     105                     110
Met Ser Thr Glu Phe Phe Val Val Pro Thr Leu Ser Ala Ser Ser Ser
        115                     120                     125
Ser Ser Thr Thr Ser Thr Thr Thr Pro Ala Thr Pro Arg Arg Ser Ser
    130                     135                     140
Gly Ala Lys Ala Arg Ala Ala Gly Thr Thr Gly Arg Val Thr Pro Gly
145                     150                     155                     160
Ser Ser Ser Asn Ser Thr Pro Pro Ser Ile Val Ser Thr His Leu Pro
                165                     170                     175
Phe Ala Gly Leu Arg Lys Glu Pro Trp Val Val Pro Val Leu Val Leu
            180                     185                     190
Ala Thr Leu Thr Met Leu Met Met Ala Ala Phe Glu Ile Phe Val Leu
        195                     200                     205
Phe Lys Ala Trp Arg Thr Ser Pro Ser Arg Arg His Leu Phe Leu Gly
    210                     215                     220
Gln Met Leu Leu Leu Gly Leu Phe Ala Cys Ala Ser Leu Gly Ala Ile
225                     230                     235                     240
Ile Thr Ala Gln Pro Ser Leu Ile Ser Cys Gly Ala Ile Arg Phe Gly
                245                     250                     255
Val Gly Val Ala Tyr Ala Leu Val Phe Ala Ala Leu Leu Val Lys Cys
            260                     265                     270
Val Phe Leu Ile Ser Leu Asn Gly Gly Val Tyr Leu Pro Ala Pro Tyr
        275                     280                     285
Gln Gly Leu Leu Leu Leu Phe Ala Leu Leu Ile Gln Val Ala Ile Gly
    290                     295                     300
Gly Gln Trp Leu Leu Thr Gln Pro Pro Glu Val Tyr Thr Thr Ser Val
305                     310                     315                     320
Pro Val Met Gly Ser Gly Phe Leu Ser Thr Thr Val Ala Ser Gln Thr
                325                     330                     335
Asn Tyr Ser Ala Leu Phe Tyr Pro Thr Ser Tyr Thr Thr Leu Asp Gly
            340                     345                     350
Thr Pro Glu Ile Tyr Thr Arg Ile Ala Ala Val Ser Thr Val Leu Ile
        355                     360                     365
Pro Leu Cys Lys Thr Gln Phe Ser Glu Leu Leu Phe Ser Leu Ile Tyr
    370                     375                     380
Ile Val Phe Leu Ile Val Phe Ile Ala Val Leu Ala Ile Lys Ser Arg
385                     390                     395                     400
Gly Ile Arg Asp Asn Tyr Arg Glu Ala Thr Tyr Ile Gly Leu Ala Ile
                405                     410                     415
Gly Gly Ala Ile Pro Ile Trp Leu Gly Trp Met Leu Cys Gly Leu Ala
            420                     425                     430
Val Ala Glu Arg His Lys Asp Ala Cys Val Ala Phe Gly Leu Val Ala
        435                     440                     445
Thr Ser Ala Thr Val Phe Leu Val Met Phe Met Pro Lys Gly Arg Gln
    450                     455                     460
Leu Ala Ala Met Gly Lys Glu Gly Leu Tyr Val Glu Asp Arg Glu Glu
465                     470                     475                     480
Gln Phe Ser Ser Leu Ser Arg Ala Gly Ser Gly Tyr Ser Pro Ser Phe
                485                     490                     495
Phe His Phe Lys Pro Ile Lys Tyr Gly Val Met Ser Gly Cys Gly Leu
            500                     505                     510
Pro Asn Ser Ala Ser Asn Thr Gly Gln Gly Leu Ser Ser Lys His Cys
        515                     520                     525
Ser Ser Ala Asn Asn Gly Gly Gly Lys Ser Arg Ser Tyr Glu Leu Phe
    530                     535                     540
Pro Ser Lys Phe Val Thr Phe Ser Ala Arg Gly Gly Arg Thr Gln Leu
545                     550                     555                     560
Lys Leu Asp Thr Val Pro Phe Arg Pro His Ser Ser Ala Leu Val Ser
```

565 570 575
Cys Gln Pro Leu Leu Ile Leu
580

<210> 79
<211> 3524
<212> DNA
<213> Drosophila

<400> 79
aaatttgtat aactttcagt tcttatactc agttattcca ttaatctcta atcccacgac 60
atttaaatac acatacaaat tttttataca aagtttgctt tattgactat ttttgaagca 120
catttaccat atctgtatta ttaaaatact tagcatagat tacagaattt tcagcttatt 180
ccatgaaata ctgtttgtca aacaaacaga tttcatcagt ttttgttttt ataatccttt 240
ttttttggga aagattctca tatatttaca attataattt aaaatagtct ctactttagc 300
ttactacata ttcaattaaa tctcaggagg ctgacttttc cgcggacttt tatggtggct 360
ctgacccaac tggatgcatc ctgcgccttt cctcgcagaa aagtatttta tttcacattt 420
tacgggcgtc ttggctggct gtcttacctc tggctaagat gcttatttta tacaaatccg 480
acttttgcga ttttaattgc ataaatcagc ttagttattt aattggttta ttaatctacg 540
ggagaagcaa ccattttatt ttaggtaagc actaagaaaa agattggatc ttaagattcc 600
gatattatta tatatcttta tttacttagt aaaaggatgg ggatttacgt atattttatt 660
taagagggta tattggcagt aaaatattta ttcaaattgt gggttgtaca gtttatatta 720
attattaata tgttattagc tgtcaaggca ataagaaaat gtaatccatt tcagcccccct 780
tcagttaaat atattgctaa catatcgatg tgcggttgta cagactcatt ttcaaatctt 840
tgaattttaa ttctgatctc ggtttttatt acgtagacac atattaatac tgcttagtct 900
taagttcttc attcatattt ggataagaaa ggattataaa atcatttaaa tgcatactat 960
aatattggct tatttaaagg aaattcctac ttaaactggg ttacaaatct gtctgctgca 1020
ctgcgcccga tcctcttctc gcggtggtgc aaaatgcctc ctgctgctcg ctcaaacata 1080
taatcaaatt ctccggcgga gctgtcactt tgccattgga attgcatgat ttctgtgtcc 1140
ccgtggttcc ggtcttttcc gctgtcgtgg tcgaggatgt ggtggtcgat gaattgttgt 1200
tacccgcatt cagtggcagc agtggaggaa tctgcggatt gctggtcaac aggcgcttgg 1260
agccacccct ctttttactg ccctgtttgg tgaaaacact gggctccagc tggagttgag 1320
cgttgatatc ctgcttattc atacaggtaa aggccttgaa gaagctcttc cggaagttct 1380
cacttaggaa ggcataaagt atgggattca ccgccgaatt cgagtaaacc agtgccccca 1440
gaagtaggaa aatgagtatt tccagtcggg agaggtccct ttgcgcggga ttcgagtgaa 1500
tcagggccac cttggagatc cagtggtgtg taaataaaaa atcgaatata ctttataatt 1560
aattacctga gaaatccagt gagggagcca acatagaatg tatacactta tcacctgcgg 1620
attgaaaaga aggtgggaat ctacaaatga attgaactga ttatttcggc attattaccg 1680
tcagtaccag tcgagtgacc ttcctgtgag cccgcctctt ctccttggac ttcgttcctg 1740
gtttgggacc caccgatcgc agtttcctta taaccaagta gtagaaactc aggataaagc 1800
acagcggtgt ggcgaatcct aggaaaaatg tgtacagtat gaaggtggtg cccgaatgct 1860
tcttgtacgc atctggccac attatgttgc acgagtaatt gatgccatcc tcctgctcca 1920
cagtgctggc ataaaggatc acgggcagca tgaggaccgc tgaagttgac caggcaatcg 1980
ctgagaccac tttggcaata tgcagagttc gatatcgtgg cgaggaaatc gggtggcata 2040
ccgctatata tcgatccgcg gacatgatga gcagaaaaat cgacgaggtg aaggaggtga 2100
tggatgtgct caccatgtag gctttgcaca taaactcccc gaatcgccag ctgcaaattc 2160
gcattgtgta cagcagaaag ggtattccaa tcaggaagca ctcgtctgcc accgccagat 2220
tcaggatata tatattcgtg accgtttgca ttttggaaaa gcgcaacacc acgtagatca 2280
ccagggtgtt gccaaataat ccgataatgc acacaaatcc gtagagcacc acagtgaaca 2340
agtcagcaaa ggaattccgc gtggctatgc aatgttgata ggtgggcaaa tcagtgccat 2400
aaagggactc ctctggctga attgtggaac taccactaat ccatctatgg ttcagttccg 2460
tggtatataa gctctcgttc gtgttatacc aatttgtggc atttgcctct gaatcggctg 2520
tcatttcccc tttcacaaac tggaggacat ccatcatcag ccacgtaaac atatttccac 2580
ctaagcgctg tccagttcaa atttaattat tttaaatagg ttttacgcgt tgttgctccc 2640
cgctcgtttt gttcttgcat tttggcgttg ctgcaaaata cattttgtta ttacgtgaaa 2700
tattttcaga cacagttgtt ttggccaaac ttgattagtt actaattgaa atactgtggt 2760
gctattaaag aaaaaaacct aaaaccttca ctagtctagt caatattttt tattccaacg 2820
taagtttaaa atgtactgta acaattggtg ataagtattc aaatggaatt aagttaagtc 2880
tacaagtttt ataagatatt ttaatataac tagattgcag ctctttaatg catagtgtat 2940

```
atttttaata tattatgagt ttacttttaa gtggcagtca attttatttg ctgtcttaac 3000
atctttgaat cgtatgtacc ccattcgtaa tgccacttca catacagatt cttttgcaac 3060
tttactacaa cttttactga ctttctgaga aatcccctcg aaaattgaaa actttcctga 3120
acaacttgta gaaaaacaaa cctttaaaaa tacgacgtat gaaataaatc gtctgtttcc 3180
tgcttttgaa aagtgagtta ccacaaattt ggtttgtaag gcgaaggaaa ttaacattta 3240
tatttaaatt aaattgctca cagcttgaca cgtccaacaa acaagaaaca agaacttttt 3300
tatattctga cctaactcga cactttctgt aaataaatta ggcgtaagga aaagctgggc 3360
agcatttatg caaatatcag gtgaaaatat ttgtcttctc ggaagtcgga aaactccagc 3420
tatacagatc cccatggttt tgtattagga gattgaaatt taagccgatt tgtttgtgca 3480
ttccatgcac atcaaatatt gttgactgag ccaaaacaca tgtg                  3524
```

<210> 80
<211> 1356
<212> DNA
<213> Drosophila

<400> 80

```
atgacagccg attcagaggc aaatgccaca aattggtata acacgaacga gagcttatat 60
accacggaac tgaaccatag atggattagt ggtagttcca caattcagcc agaggagtcc 120
ctttatggca ctgatttgcc cacctatcaa cattgcatag ccacgcggaa ttcctttgct 180
gacttgttca ctgtggtgct ctacggattt gtgtgcatta tcggattatt tggcaacacc 240
ctggtgatct acgtggtgtt gcgcttttcc aaaatgcaaa cggtcacgaa tatatatatc 300
ctgaatctgg cggtggcaga cgagtgcttc ctgattggaa taccctttct gctgtacaca 360
atgcgaattt gcagctggcg attcggggag tttatgtgca aagcctacat ggtgagcaca 420
tccatcacct ccttcacctc gtcgattttt ctgctcatca tgtccgcgga tcgatatata 480
gcggtatgcc acccgatttc ctcgccacga tatcgaactc tgcatattgc caaagtggtc 540
tcagcgattg cctggtcaac ttcagcggtc ctcatgctgc ccgtgatcct ttatgccagc 600
actgtggagc aggaggatgg catcaattac tcgtgcaaca taatgtggcc agatgcgtac 660
aagaagcatt cgggcaccac cttcatactg tacacatttt tcctaggatt cgccacaccg 720
ctgtgcttta tcctgagttt ctactacttg gttataagga aactgcgatc ggtgggtccc 780
aaaccaggaa cgaagtccaa ggagaagagg cgggctcaca ggaaggtcac tcgactggta 840
ctgacggtgg ccctgattca ctcgaatccc gcgcaaaggg acctctcccg actggaaata 900
ctcatttccc tacttctggg ggcactggtt tactcgaatt cggcggtgaa tcccatactt 960
tatgccttcc taagtgagaa cttccggaag agcttcttca aggcctttac ctgtatgaat 1020
aagcaggata tcaacgctca actccagctg gagcccagtg ttttcaccaa acagggcagt 1080
aaaaagaggg gtggctccaa gcgcctgttg accagcaatc cgcagattcc tccactgctg 1140
ccactgaatg cgggtaacaa caattcatcg accaccacat cctcgaccac gacagcggaa 1200
aagaccggaa ccacggggac acagaaatca tgcaattcca atggcaaagt gacagctccg 1260
ccggagaatt tgattatatg tttgagcgag cagcaggagg cattttgcac caccgcgaga 1320
agaggatcgg gcgcagtgca gcagacagat ttgtaa                           1356
```

<210> 81
<211> 451
<212> PRT
<213> Drosophila

<400> 81

```
Met Thr Ala Asp Ser Glu Ala Asn Ala Thr Asn Trp Tyr Asn Thr Asn
1               5                   10                  15
Glu Ser Leu Tyr Thr Thr Glu Leu Asn His Arg Trp Ile Ser Gly Ser
            20                  25                  30
Ser Thr Ile Gln Pro Glu Glu Ser Leu Tyr Gly Thr Asp Leu Pro Thr
        35                  40                  45
Tyr Gln His Cys Ile Ala Thr Arg Asn Ser Phe Ala Asp Leu Phe Thr
    50                  55                  60
Val Val Leu Tyr Gly Phe Val Cys Ile Ile Gly Leu Phe Gly Asn Thr
65                  70                  75                  80
Leu Val Ile Tyr Val Val Leu Arg Phe Ser Lys Met Gln Thr Val Thr
                85                  90                  95
Asn Ile Tyr Ile Leu Asn Leu Ala Val Ala Asp Glu Cys Phe Leu Ile
```

```
                100                     105                     110
Gly Ile Pro Phe Leu Leu Tyr Thr Met Arg Ile Cys Ser Trp Arg Phe
            115                     120                     125
Gly Glu Phe Met Cys Lys Ala Tyr Met Val Ser Thr Ser Ile Thr Ser
        130                     135                     140
Phe Thr Ser Ser Ile Phe Leu Leu Ile Met Ser Ala Asp Arg Tyr Ile
145                     150                     155                     160
Ala Val Cys His Pro Ile Ser Ser Pro Arg Tyr Arg Thr Leu His Ile
                    165                     170                     175
Ala Lys Val Val Ser Ala Ile Ala Trp Ser Thr Ser Ala Val Leu Met
                180                     185                     190
Leu Pro Val Ile Leu Tyr Ala Ser Thr Val Glu Gln Glu Asp Gly Ile
            195                     200                     205
Asn Tyr Ser Cys Asn Ile Met Trp Pro Asp Ala Tyr Lys Lys His Ser
        210                     215                     220
Gly Thr Thr Phe Ile Leu Tyr Thr Phe Phe Leu Gly Phe Ala Thr Pro
225                     230                     235                     240
Leu Cys Phe Ile Leu Ser Phe Tyr Tyr Leu Val Ile Arg Lys Leu Arg
                    245                     250                     255
Ser Val Gly Pro Lys Pro Gly Thr Lys Ser Lys Glu Lys Arg Arg Ala
                260                     265                     270
His Arg Lys Val Thr Arg Leu Val Leu Thr Val Ala Leu Ile His Ser
            275                     280                     285
Asn Pro Ala Gln Arg Asp Leu Ser Arg Leu Glu Ile Leu Ile Phe Leu
        290                     295                     300
Leu Leu Gly Ala Leu Val Tyr Ser Asn Ser Ala Val Asn Pro Ile Leu
305                     310                     315                     320
Tyr Ala Phe Leu Ser Glu Asn Phe Arg Lys Ser Phe Phe Lys Ala Phe
                    325                     330                     335
Thr Cys Met Asn Lys Gln Asp Ile Asn Ala Gln Leu Gln Leu Glu Pro
                340                     345                     350
Ser Val Phe Thr Lys Gln Gly Ser Lys Lys Arg Gly Gly Ser Lys Arg
            355                     360                     365
Leu Leu Thr Ser Asn Pro Gln Ile Pro Pro Leu Leu Pro Leu Asn Ala
        370                     375                     380
Gly Asn Asn Asn Ser Ser Thr Thr Thr Ser Ser Thr Thr Thr Ala Glu
385                     390                     395                     400
Lys Thr Gly Thr Thr Gly Thr Gln Lys Ser Cys Asn Ser Asn Gly Lys
                    405                     410                     415
Val Thr Ala Pro Pro Glu Asn Leu Ile Ile Cys Leu Ser Glu Gln Gln
                420                     425                     430
Glu Ala Phe Cys Thr Thr Ala Arg Arg Gly Ser Gly Ala Val Gln Gln
            435                     440                     445
Thr Asp Leu
        450
```

<210> 82
<211> 7448
<212> DNA
<213> Drosophila

<400> 82

```
agcatgcggc tgaacaaaaa aaaactcgag gaatggccgc aatttcggtg agccatataa 60
acaacataac atatagtttt cccatccatc tacgctgtct gcgccactgg ccaagtccaa 120
gttgtcttga cggcccggca aaaagcaaca tccaaggaaa atgttgccat tttagatgcg 180
ccgccccgctg agcggaaaag tcgcgcggaa attactaaat cgatacgtgc tgcgtatcag 240
attttccatg ggttctgcat gcaaaaatac cttgaaatcc ttgcctgcag tcgcacacat 300
acacgtattg ccatctttgt acagacaagc ttgactaagt cgaactttac tatgttattc 360
tactaaaagg ttttaaagca aatgatctat actatactat actattaact tcaatagaat 420
```

```
tcctattttc tctggctagc aaacaaattt ggatcataaa aagattttca tggcatttat 480
gagatacgta gattgtgtgg ctcaaccttt tatctatatt atttccataa atttaagcac 540
caatggaaat gtttatctcc aaagcttagg aaaatttgga agtatatata ataaaaacct 600
tttatgtctg caaataagag atattacgtg tatactgagt ttcgcttgag tttcggagtt 660
aagaggcttc cctatagact gtttgaaaat cctataatgc cttggcattt ctaaaaggtc 720
aactgaaagc agtggaattg ctggtgggcg aggttcgtct cctaagtggc aaaatgatcc 780
agaatcggaa gcaaaccctg acgatgaaat aaactatacg agaaactgga gaacaattac 840
acatggactg tgggtgcaat tagggaccac accttttttcg ctgccgacgc gtctaaagtt 900
aatgaatatg ttaataatat gttatgcgcg tgcaaatttg ttcagggctt gaggggtttt 960
ctgtgtttct gtttccaggc gagtccagtc caatccaact ctatcgcctg atggacatgg 1020
tgttcgtggg gaactgtggc ttctggccgg cgctgaagaa gcgtttgcag gtgaggcgcc 1080
agaaggccgc ccggaagtcg acgctgtaga aggcgtagat gaagggatta atggcgctat 1140
tgaaccaacc cagccaggtg agggccttgg ccagcatttg gctggcctgg tgttcggcca 1200
ggaacggggt gatcagatag ttgataaaga agggcagcca gcaggcgata aagccaccca 1260
ccacaatgct cagggtttga gtggtcttgt tctccttctt cagcgacgta atccgattgg 1320
acagcgattt actgttggtg ttggcactgg tgttcccgct agtcgatgtc gtggttacgc 1380
ccgcctggtg atgataatga ggatggtggt gatggcttcg cagtctctcg cggtcccgct 1440
ccccatgcga atgcctaaag ctgtgtgttc gcggtggctg cgatagggag tgggtctgca 1500
cctggcccgt ttggcccggc ggctgggact gccattgcgc atccaggaga ctactgcctg 1560
ggcccatgcc actcgtcatt gtggtaatgg tcgttgaggc ggtggaagcg cgcttcagcg 1620
aggagggtcg tgtgagttcg tagcacccat tcttggcgca gcccatggag ccacctactc 1680
cgccggaggg aagggccagc agggattgac agtgtgtgcc gccagtggca ctacttgctc 1740
ctccaccccc accacctgct cctccagcgc ccatcgccgc acaattgtcg ctgtcgctga 1800
gcatcatgtc ttggagctcc ttggctatcg tctggttgga gaacgttcgc gatgtggcct 1860
gtctctggcg ctggcccacc gagctgtgct gctcctgctc cgacagctca ttctccacgt 1920
ccgctgccgt gtagcgcttg aatttctaga cacgcaaagt gcacaaggcg gttaagtcaa 1980
tgtcgatgat aaataactaa aatagagtca atgggcggca aaaggcggca aagataaaag 2040
aaacggaaac ggaagtcacg ctaggaaaag ccgaagcttt aaacactctt gtaggcatca 2100
tcatatactt taaagggaac cacactggta agcagatggc tttaagctga tttaagacta 2160
catacccaaa tgccatgcgg ccattgttct atcaactgca cgatgtgcta atctctctaa 2220
acataaaact gaatttcaga ccggttttaa tgcgtatgga gctgcctgga cttgattcag 2280
tttaagtgaa ttttaagcct attttaaaat gcacaccagg tacacataca aatgtatata 2340
tgcagatgca attttcgaag cgagcatcgt ttagtgttag ctatttagtg atgctaactt 2400
aaagctcgac ataaatgagt cgaatatgta attaccctgg tattcaagga tgtagcaaca 2460
ttagcaacat ttcataacta acatttttga aaagtattca actctaaagc gccaaactgc 2520
tctctcgtta gaaaacggtt agttgaaaac tcattaattc agctatacgg gctgccagag 2580
ttgaagtgaa aaacttgttt tatagcggac actggaccgg aaattggcac ataccccgat 2640
gcgagggtat ggtatttaaa tctgtgtgag ctcctggcat ttgggaatta tgaggcaaac 2700
gtgccacagt ttgtgccgca cgtggagtgt cctttaagcc ggagtggttc gcaatagaga 2760
tgactttcgc tggtcgggcc agcgtttgtt gtttgcagct gccgtggttt tcctattcca 2820
tattccgtat ttgtgcacta tgttggcttt ggggttttgg atggcgggag acgtgcgaat 2880
gtttgcaggc aaacagaagg acatgcggac attgggacat ggttgcggat ggctgcgacg 2940
gagcttacct tgttgtgaac acttatgtcg gtcatattgt cgtgcctgga tgcaatgaca 3000
caggaaattc ttgcatacac ataaatcatg actgccatgg gtatgaagaa ggagcccatg 3060
gccgagaaga tgacgtagcc ctcgttctgg ttgtatcggc actcgcgcag atcccttcgt 3120
cctggctcgt acctgcggga aatggaagtg agcaggatac gcttaatggg ccaccaagtt 3180
agtcaactaa tgcaggaagt tcccaattat ttgactacgg agctgtgacc ctcggagggg 3240
tcagaggtcg caagagcagc agacaattgc cagctaatta tattcgtttg ttcagctggg 3300
tgtcgccagt cagtctgctc ggaggtctcg aaatttattc taattaattt acactcacac 3360
tcggcgacat ttgtaatgct taaaatggac caaaaagata cccaaagtgg tccttcaaaa 3420
ttatgcaatg gcatcttttg acacaatttt ctgcagtcaa caaaattgtt gatttaattt 3480
catttaagca aattaagctt taattatttt tgtaagttaa ataaattaaa catctcttgg 3540
ttacatttcg gaattctgct gcctacagga acatttagta tatttcttca aaaaatctgt 3600
tgactgtgct aacgaagtta gatataaaca aagcacttta attaattctc cttttagtta 3660
gttagacaaa agtagagaga acaaaatttg ttcttaaagc tagcagatgc ggtataccaa 3720
caaggaagtt gatgtttagg ttaaaattcg gtgcatctat ttagtgtgtg gactgttgac 3780
caacttagaa gacttcttgc tgcactcctt cttcgttatg gaaataaccc tacatactga 3840
ttatttcttc aagtgtactt accatcccag catgggggga caggtgatgg ccaacgcgag 3900
cagccagacg attaggatca tgatgagggc caatcgtttg gagcgccgtt tacgggagta 3960
cgtgagcggt ctggtcacgg ccaaatatct gtcggatgga cgtgaaggat gcaaagattt 4020
```

acgatatcag gggaatacta tggatgtttt ggataggcct atgctcacct gtccacactg 4080
atggcgcaca ggctgagaat ggatgccgtg cagaggagca cgtccaggga gatccaaatg 4140
tcgcagagca cccagcccag ttgccacgag cctgcaatca agattccaaa ttaaaaactt 4200
tccgccagcc ttttgcagcg cggagtgggg cagccaactt aaaattccgg tccaaagtca 4260
atttgattag caacattttc catggctttc agcaaatgcc tcacctaagg ccaattaaag 4320
cggccagctc ggcaaatttt ggcaatgttt caggtcccct tcaacccgtt agtggtttaa 4380
ttctataccg aaaaagcacc gatttctcca tttgacccca acacttcgag cttctacagt 4440
gctttctcct cgtcggtggc ttgtcaagcg attaattttt aaatcaagtg ccccttaatg 4500
gtctcccggc gccaaatgaa aagttgtttt tgtgaaaggc gctaggagcg aaggaaccaa 4560
ggagctcatt acacggatat tgcctatggg ccagcacact tgggttggtc tagttcaatg 4620
taaggttgct gaatttgctg aatctccaag gactcggatt gaagcagtta aaggcaaaat 4680
ttggagtgaa ttagaaaaac acaaggatta cttaattaat tatgggaatg cttgttcagt 4740
tgtagtgtat gtgtatgttt ttatctaaaa actggagtag gaactggcac aaagcagtcg 4800
tcatatcagc actttcagtg cctaccccaa tttcaactgg atgataaggc tccaaaaagc 4860
aaactataaa tcccgaccgt gtgggtgtat tacgtcctaa cttaatgata tatgtgcgta 4920
tactttggac aaaaatgaac tgaatgtggg gcttggtaaa ttgtttgcct cgaaaaaacc 4980
tgggggcatg agtaattatg aatgagattc aaaaggtatc tccacagaga gaaactagtg 5040
gatgaaaatc ggaaaacaag cttaatgttc cataagacag aagccaagtg agctcaaatc 5100
agaagggttt agcatatttt ccattcaact tcacgagttc agcaacaaag atcaactatt 5160
ttctgacgag aataattcat tccgattcgg attgtcagtg gtccggatga ttatgacaaa 5220
ctttcgacat gccgttgcct ttaaagtatt tccacaaagt tgggctcggg aattttcccc 5280
tttttcggct tttttttttcg tgctgaggaa attgcacgag agccgaggca aaaaaatgaa 5340
agaaattctt tcagttcaaa gaatacaaat ccataaggcg agagagatac atactcgcat 5400
atttcaccta gcagcattta cataaataca cgccccccttt gccccgcccc accatctgac 5460
acacctcctc ctcgagttcc tgttcgtgtc cgtgctgcta ttcctgttcg tgctactttg 5520
ttgggataca cttggagaaa aatatatata tatgctagtc ctcctagaaa tccctacaca 5580
ttgctctaaa cagaatctca ttgatattta caggttcgaa agttccttaa atacctcaag 5640
aaacattgct actacatttt tacttcatga gatacaaatt acaaagactt gtattggatt 5700
tgttctcagt gcattcgctt ggttctgggc gccaagcttt tgttagcaat tttttgaggg 5760
atttaccgga atggcatctt catcttgtaa tcacgcgaaa agcccggcta gaggcgcctc 5820
ctttcgcgga aggactctaa tgccttaatg gcgctgccac acactcacac acacacttac 5880
acactcacac tcatacacac atagactcac ctatgagatg gacggcgacg gcggggggca 5940
tgacgaagat gcccacgaga aggtcggcca ccgccaggct catcacaaag cagttggtaa 6000
tggtgcgaag ccggcgagtg gtcaatatgg ccagtatgac cagggtgttc ccgatgacgg 6060
tgaccacgat gaaggtggca aagactgcaa tcaggatgat gccctgccag gagagcacca 6120
gatcgtagaa gtgcgcccat tcggcggccg ccgccgactc gttgtagtaa ttaaatgtgg 6180
tcacgggcgc cgtgggtccg ggttcggtga gagccactgc atccgtaccg ttcggtccgc 6240
tgggatatcc acctcctact cctcctactc ctcctccacc tgtcccattg cccgcatcct 6300
catatacgat ggatatgttt tccggcagac caccccactcc cactcccgat cccagtccca 6360
atccccgcag ggaggaggag gaggcagcca catacatgag cgacggataa atggcgctta 6420
aattattcag ctgcaagttg acagccatgg atcggccagt tgggttcaca atcgagtgtg 6480
gctagctccg gcagcaggcc gcctaggccc atccatatcg tgccatatcc ctggagttga 6540
cggagcagga acacggccca actgtggtca ttgtgtccag ttgcacgtcg ggttaactgg 6600
aattaaagga aagttcgaca tttagcaaag tcctttgcat ggtaataaaa acgaagtcaa 6660
gaggaaacaa taaccatttg catatccctt ttatgctgaa atctatttgt gtttgtttta 6720
attttttatg tatgcacgca ttatctgatt tattccttgt gacaatttcc gaaattaaat 6780
tgatttcaca gcattccata caatctcatt taaattgcaa tcgtcgaaat tgtcaatttc 6840
attcattcgc ttgcaattaa gcattcaaca atctcaaggc aaaaagtgca cagcattgaa 6900
tatacccgtc tgtgtaaatt acgttgaaag ttgtcaacgg tgtcaaaaga ttcttaagat 6960
ggcctttgta ttgggttgcc attaccgcga tacagttaaa tactttactt tctaagtaga 7020
taaacgttaa tcctttagtt agtcatgctt tcaacgaaaa tgccgaagtg cgctactgtg 7080
agcaataaat caaattaagt ttaccttatc atctagaagc acccactggc ttgatttaat 7140
ttaaaatgta tgattccatt tgctttggat taggtaccga acttagagac ttaagaataa 7200
ataattcgtg gaggtatttt ctatcagaaa attaaaaatt tttatacaaa tattcacggt 7260
ttgatttggt gaagaaaatt caagaaaatc atcaacttcg gaattcctgg cgcattctaa 7320
taatcattat atgaaaatta caaatataat gagctttata caaagtaagc actattacta 7380
caaagtggcg aaatgtatta ccaataaatt tgccaaggga atggctacat tttgggaata 7440
ttttattt 7448

<210> 83
<211> 1896
<212> DNA
<213> Drosophila

<400> 83
atggctgtca acttgcagct gaataattta agcgccattt atccgtcgct catgtatgtg 60
gctgcctcct cctcctccct gcggggattg ggactgggat cgggagtggg agtgggtggt 120
ctgccggaaa acatatccat cgtatatgag gatgcgggca atgggacagg tggaggagga 180
gtaggaggag taggaggtgg atatcccagc ggaccgaacg gtacggatgc agtggctctc 240
accgaacccg gacccacggc gcccgtgacc acatttaatt actacaacga gtcggcggcg 300
gccgccgaat gggcgcactt ctacgatctg gtgctctcct ggcagggcat catcctgatt 360
gcagtctttg ccaccttcat cgtggtcacc gtcatcggga acaccctggt catactggcc 420
atattgacca ctcgccggct tcgcaccatt accaactgct ttgtgatgag cctggcggtg 480
gccgaccttc tcgtgggcat cttcgtcatg ccccccgccg tcgccgtcca tctcataggc 540
tcgtggcaac tgggctgggt gctctgcgac atttggatct ccctggacgt gctcctctgc 600
acggcatcca ttctcagcct gtgcgccatc agtgtggaca gatatttggc cgtgaccaga 660
ccgctcacgt actcccgtaa acggcgctcc aaacgattgg ccctcatcat gatcctaatc 720
gtctggctgc tcgcgttggc catcacctgt ccccccatgc tgggatggta cgagccagga 780
cgaagggatc tgcgcgagtg ccgatacaac cagaacgagg gctacgtcat cttctcggcc 840
atgggctcct tcttcatacc catggcagtc atgatttatg tgtatgcaag aatttcctgt 900
gtcattgcat ccaggcacga caatatgacc gacataagtg ttcacaacaa gaaattcaag 960
cgctacacgg cagcggacgt ggagaatgag ctgtcggagc aggagcagca cagctcggtg 1020
ggccagcgcc agagacaggc cacatcgcga acgttctcca accagacgat agccaaggag 1080
ctccaagaca tgatgctcag cgacagcgac aattgtgcgg cgatgggcgc tggaggagca 1140
ggtggtgggg gtggaggagc aagtagtgcc actggcggca cacactgtca atccctgctg 1200
gcccttccct ccggcggagt aggtggctcc atgggctgcg ccaagaatgg gtgctacgaa 1260
ctcacacgac cctcctcgct gaagcgcgct tccaccgcct caacgaccat taccacaatg 1320
acgagtggca tgggcccagg cagtagtctc ctggatgcgc aatggcagtc ccagccgccg 1380
ggccaaacgg gccaggtgca gacccactcc ctatcgcagc caccgcgaac acacagcttt 1440
aggcattcgc atggggagcg ggaccgcgag agactgcgaa gccatcacca ccatcctcat 1500
tatcatcacc aggcgggcgt aaccacgaca tcgactagcg ggaacaccag tgccaacacc 1560
aacagtaaat cgctgtccaa tcggattacg tcgctgaaga aggagaacaa gaccactcaa 1620
accctgagca ttgtggtggg tggctttatc gcctgctggc tgcccttctt tatcaactat 1680
ctgatcaccc cgttcctggc cgaacaccag gccagccaaa tgctggccaa ggccctcacc 1740
tggctgggtt ggttcaatag cgccattaat cccttcatct acgccttcta cagcgtcgac 1800
ttccgggcgg ccttctggcg cctcacctgc aaacgcttct tcagcgccgg ccagaagcca 1860
cagttcccca cgaacaccat gtccatcagg cgatag 1896

<210> 84
<211> 631
<212> PRT
<213> Drosophila

<400> 84
Met Ala Val Asn Leu Gln Leu Asn Asn Leu Ser Ala Ile Tyr Pro Ser
1 5 10 15
Leu Met Tyr Val Ala Ala Ser Ser Ser Ser Leu Arg Gly Leu Gly Leu
20 25 30
Gly Ser Gly Val Gly Val Gly Gly Leu Pro Glu Asn Ile Ser Ile Val
35 40 45
Tyr Glu Asp Ala Gly Asn Gly Thr Gly Gly Gly Gly Val Gly Gly Val
50 55 60
Gly Gly Gly Tyr Pro Ser Gly Pro Asn Gly Thr Asp Ala Val Ala Leu
65 70 75 80
Thr Glu Pro Gly Pro Thr Ala Pro Val Thr Thr Phe Asn Tyr Tyr Asn
85 90 95
Glu Ser Ala Ala Ala Ala Glu Trp Ala His Phe Tyr Asp Leu Val Leu
100 105 110

```
Ser Trp Gln Gly Ile Ile Leu Ile Ala Val Phe Ala Thr Phe Ile Val
        115                 120                 125
Val Thr Val Ile Gly Asn Thr Leu Val Ile Leu Ala Ile Leu Thr Thr
    130                 135                 140
Arg Arg Leu Arg Thr Ile Thr Asn Cys Phe Val Met Ser Leu Ala Val
145                 150                 155                 160
Ala Asp Leu Leu Val Gly Ile Phe Val Met Pro Pro Ala Val Ala Val
                165                 170                 175
His Leu Ile Gly Ser Trp Gln Leu Gly Trp Val Leu Cys Asp Ile Trp
            180                 185                 190
Ile Ser Leu Asp Val Leu Leu Cys Thr Ala Ser Ile Leu Ser Leu Cys
        195                 200                 205
Ala Ile Ser Val Asp Arg Tyr Leu Ala Val Thr Arg Pro Leu Thr Tyr
    210                 215                 220
Ser Arg Lys Arg Arg Ser Lys Arg Leu Ala Leu Ile Met Ile Leu Ile
225                 230                 235                 240
Val Trp Leu Leu Ala Leu Ala Ile Thr Cys Pro Pro Met Leu Gly Trp
                245                 250                 255
Tyr Glu Pro Gly Arg Arg Asp Leu Arg Glu Cys Arg Tyr Asn Gln Asn
            260                 265                 270
Glu Gly Tyr Val Ile Phe Ser Ala Met Gly Ser Phe Phe Ile Pro Met
        275                 280                 285
Ala Val Met Ile Tyr Val Tyr Ala Arg Ile Ser Cys Val Ile Ala Ser
    290                 295                 300
Arg His Asp Asn Met Thr Asp Ile Ser Val His Asn Lys Lys Phe Lys
305                 310                 315                 320
Arg Tyr Thr Ala Ala Asp Val Glu Asn Glu Leu Ser Glu Gln Glu Gln
                325                 330                 335
His Ser Ser Val Gly Gln Arg Gln Arg Gln Ala Thr Ser Arg Thr Phe
            340                 345                 350
Ser Asn Gln Thr Ile Ala Lys Glu Leu Gln Asp Met Met Leu Ser Asp
        355                 360                 365
Ser Asp Asn Cys Ala Ala Met Gly Ala Gly Gly Ala Gly Gly Gly Gly
    370                 375                 380
Gly Gly Ala Ser Ser Ala Thr Gly Gly Thr His Cys Gln Ser Leu Leu
385                 390                 395                 400
Ala Leu Pro Ser Gly Gly Val Gly Gly Ser Met Gly Cys Ala Lys Asn
                405                 410                 415
Gly Cys Tyr Glu Leu Thr Arg Pro Ser Ser Leu Lys Arg Ala Ser Thr
            420                 425                 430
Ala Ser Thr Thr Ile Thr Thr Met Thr Ser Gly Met Gly Pro Gly Ser
        435                 440                 445
Ser Leu Leu Asp Ala Gln Trp Gln Ser Gln Pro Pro Gly Gln Thr Gly
    450                 455                 460
Gln Val Gln Thr His Ser Leu Ser Gln Pro Pro Arg Thr His Ser Phe
465                 470                 475                 480
Arg His Ser His Gly Glu Arg Asp Arg Glu Arg Leu Arg Ser His His
                485                 490                 495
His His Pro His Tyr His His Gln Ala Gly Val Thr Thr Thr Ser Thr
            500                 505                 510
Ser Gly Asn Thr Ser Ala Asn Thr Asn Ser Lys Ser Leu Ser Asn Arg
        515                 520                 525
Ile Thr Ser Leu Lys Lys Glu Asn Lys Thr Thr Gln Thr Leu Ser Ile
    530                 535                 540
Val Val Gly Gly Phe Ile Ala Cys Trp Leu Pro Phe Phe Ile Asn Tyr
545                 550                 555                 560
Leu Ile Thr Pro Phe Leu Ala Glu His Gln Ala Ser Gln Met Leu Ala
                565                 570                 575
Lys Ala Leu Thr Trp Leu Gly Trp Phe Asn Ser Ala Ile Asn Pro Phe
            580                 585                 590
```

Ile Tyr Ala Phe Tyr Ser Val Asp Phe Arg Ala Ala Phe Trp Arg Leu
595 600 605
Thr Cys Lys Arg Phe Phe Ser Ala Gly Gln Lys Pro Gln Phe Pro Thr
610 615 620
Asn Thr Met Ser Ile Arg Arg
625 630

<210> 85
<211> 3836
<212> DNA
<213> Drosophila

<400> 85
gtgctcgaca tttcgcttgc gtcccgtaaa gaaattgtcc accacgatga cctcgtgccc 60
ctggaccatc agatcgtcga ccaagtggga gcccacgaag ccagctcctc cggtaattag 120
tatgcgtttg cgattcttat agttcaggta tttaaccttc ggatacttgc gcggtgtgct 180
ggtctggagg ctacgaactt gttcctccaa tctggccaaa ttctcccttg tgcgctggag 240
ttcggccttt tgttctcgaa taagagagct ctgctcctcg tacgccatct gaaggctcct 300
ttgtagcgga ctttccgtgg gcggccattt ggcctgaacc tcctcgacgc cgggaacgga 360
aacggcgacc tttccgcttg ggcagaagct ggccattcgg taaaggtaga ccaggagcag 420
gaggagcagc gagatggctg caacgatctt cagacgcttt ttggtggcag tcatcgtcgg 480
ttatgggttg tgtactgcag ggttccagtg cttatcagca actgatatcg ccacattgca 540
ttctcgttcg gctttagaaa ctatataaac aatacaattg cgttatcagg aggtgccaga 600
gtggccagat ggctcaaacg gaatttctaa atggattaat aatctaacgc cagggaaaag 660
ctcccgccta attttaaaat ttacatcaaa agtatcgaca ggcgaacatt ttgattttgg 720
attataaatt ataatgaatt gataatgaat tagacatttt tgttggtgat cttccccagg 780
gacgaatatt tttcgtttta acgataaaag gtcggaaatt gccaaaaaat gagtttgaaa 840
actaaatttg tttgagggct ttcttaggct ggtcagcagg gactactttc cgaataaagt 900
gtttgcccaa acggctattg aaggctattt aaactgccaa cgatgagcat cgatcgccag 960
tcgggcaaag ttcttagtgc caagtgttgt agtgctcagc atgaggcttc cttgggtgat 1020
tttctgtaca gttctgcttt taattttac gaacaactca aatgcagata ttcccgggtg 1080
caactactac gacacggttg atatctcata cattgaaagg caaaacgatt cgtatttata 1140
tgatgacatc gaaattcctg ctagcttaac tggatactac gagttcaggc agtttgggga 1200
cggttcgatt acgccgattg aaaagcattt aagggcctgt gtttgcagcg tgcggccctg 1260
tattcgaatc tgctgcccag ccaagaactt tttggccaac ggaaaatgcg atgatggtct 1320
caaagaggag ctcgcccggt tcaaacccta tatatacttc acatacatgg acctacaggc 1380
acgagtacca ctcaccgata tggctattat cagagatgag ttctttgatt gtgatgaaat 1440
gatttacatc agcgacttta actatttctt agaagaggtt agcattcaaa tcttcaataa 1500
atgtggactt atagtttggt tccaggatgg aaaattttgg gttaccgttg acctcttcat 1560
ggaaaaacag gactactgct tgtaccggca caactttgat tcggattttc caaagtccat 1620
gtggataata cgacaccgct gcacaagcca catatctcct ggatccttag agagtaattc 1680
aattattatt tataacaaaa gtcatcaata tcgttttcat agctatatag aaagaaacga 1740
aaataactcc ttattggtgt gcaaactgct gactgcattc tgtgtgccct ttgcaagttt 1800
tttaattaca ataagctggt tttttatttc ttatattttc tttttcagtt ctaattataa 1860
caatgatatg ctttgtccta acaatcgcag tatatctata cattaagaag ctgcgaaatg 1920
ttactggcaa gtgcattgta tgctgtatag tttcgaggtt tatccagtgc ttgatcatga 1980
tactagatca tttaaatcta ttgaatggca tttgctctcc agctggttag ttcgaattgg 2040
agatatctct ggaagcttta aatgttatat cgttgacttc ttgcaggtta cagctcgcac 2100
tttttccgga tggcttccaa cctctggctc tccgtcatca gctaccatac gtggaaagtc 2160
ttgacgtcgc tcaatcgagt cgaccctaac tatcggttcc tgcggtacaa cgccttcgtc 2220
tggagcacag ccgcaatcat gacgggaagt atttatatag ttaatcagat ttgggaaaac 2280
gatcccagta aatggaactg gttgcctctg gtcggtttta ttcggtgctc ggtcaaaggt 2340
aaactttcaa tgcgattgct tttaaatcgt taccatggtt acgtattgct tttccatatc 2400
agattggcac ccatccgtct ggatctatat aagtggaccg tcgctggccc tgagcacttt 2460
caatgtcgcc atgttcgccc tgacagccat ttacattagg aaagtgaagg ggggtataaa 2520
taagttcaca aatgaggagg agggaaggat taactgcata aactttgaca gccagacgta 2580
agatatgata aaaccttgga ccttggtaaa ctagttatcc ttgccgtatt atccacatag 2640
ttacctacag ttcctgcggc tctccatcgt gatgggcctt acttggatat tcaatgtcat 2700
tccgtattct gcacggctcc acattttctg ggaatgggtc gggataatat ccgagtattt 2760

```
tcacagcgcg tttggaattg ttttgttcgt tctgctcgtc ctgaagcgca gcacatggac 2820
tctaatgatg gattcttaag cttgtttcta tataattcta ataattgttt acaattaaag 2880
ctttaaacaa ttaaagcccc tcattttatt aatatccttt gctgaacgaa ggaattgaat 2940
tgaattttat tgaactcaat aaataaatat agttttagaa aacgatgtac tttaaactgc 3000
caagaggtca ccgcgacaag cttttcactt cgccaccctc tggtaacact gattctccag 3060
ccactttcga ttctttaaaa aacatcgata ggtatcgcca aaaaaatcag ctgctctgct 3120
tgtagttttg cgatttcatt gcaattttaa tttgctttac tgttcaaaaa gaaatacagt 3180
gcaaaatatg ccgaacgaaa acgcagagga cctctcgggt caagagctaa agcagaaggc 3240
caaggaggtg gccgacgcat cggaggcaat gctcgaaaag gttgttgccg gcttaaatat 3300
ccaggacacg gcatcgacga atgcagcagg aaacgaggat gcggagcagc ctgatggtgc 3360
caagaatgag gcttcagtgt ctgcgaatgc aagtaagttt tcaatatgcc tgtgaaaagt 3420
catttagagg ctcctcctgc ggctgaaaaa atgaggttag gttttgtcgg gaactacagt 3480
ccaacctctc taacttgaac tgctgtgcta ttcatttcaa ttcaattttc aattttcatt 3540
ttctgtgcaa ttactccaaa gattagccct aacggaaaaa taaacaactc ataaatcata 3600
aatacatatt tgattttgat gaaatatgaa tatacttaca tagatacaat tgagtttttt 3660
attacgaata aaatatgtgt aatataagat ttctaccttg ttgattccat tgattgatgc 3720
ttcattgaaa ccccaaaatt gggcaaattt gttgtcatgg gaaactagtc cactgtcgtg 3780
atgccttaag ctttttttgg ggcagattag tcattccagt ttaaactaat cgaaaa     3836
```

<210> 86
<211> 1536
<212> DNA
<213> Drosophila

<400> 86

```
atgaggcttc cttgggtgat ttctgtacaa gttctgcttt taatttttac gaacaactca 60
aatgcagata ttcccgggtg caactactac gacacggttg atatctcata cattgaaagg 120
caaaacgatt cgtatttata tgatgacatc gaaattcctg ctagcttaac tggatactac 180
gagttcaggc agtttgggga cggttcgatt acgccgattg aaaagcattt aagggcctgt 240
gtttgcagcg tgcggccctg tattcgaatc tgctgcccag ccaagaactt tttggccaac 300
ggaaaatgcg atgatggtct caaagaggag ctcgcccggt tcaaacccta tatatacttc 360
acatacatgg acctacaggc acgagtacca ctcaccgata tggctattat cagagatgag 420
ttctttgatt gtgatgaaat gatttacatc agcgacttta actatttctt agaagaggtt 480
agcattcaaa tcttcaataa atgtggactt atagtttggt tccaggatgg aaaattttgg 540
gttaccgttg acctcttcat ggaaaaacag gactactgct tgtaccggca caactttgat 600
tcggattttc caaagtccat gtggataata cgacaccgct gcacaagcca catatctcct 660
ggatccttag agattctaat tataacaatg atatgctttg tcctaacaat cgcagtatat 720
ctatacatta agaagctgcg aaatgttact ggcaagtgca ttgtatgctg tatagtttcg 780
aggtttatcc agtgcttgat catgatacta gatcatttaa atctattgaa tggcatttgc 840
tctccagctg gttacagctc gcactttttc cggatggctt ccaacctctg gctctccgtc 900
atcagctacc atacgtggaa agtcttgacg tcgctcaatc gagtcgaccc taactatcgg 960
ttcctgcggt acaacgcctt cgtctggagc acagccgcaa tcatgacggg aagtatttat 1020
atagttaatc agatttggga aaacgatccc agtaaatgga actggttgcc tctggtcggt 1080
tttattcggt gctcggtcaa agattggcac ccatccgtct ggatctatat aagtggaccg 1140
tcgctggccc tgagcacttt caatgtcgcc atgttcgccc tgacagccat ttacattagg 1200
aaagtgaagg ggggtataaa taagttcaca aatgaggagg agggaaggat taactgcata 1260
aactttgaca gccagacgta agatatgata aaaccttgga ccttggtaaa ctagttatcc 1320
ttgccgtatt atccacatag ttacctacag ttcctgcggc tctccatcgt gatgggcctt 1380
acttggatat tcaatgtcat tccgtattct gcacggctcc acattttctg ggaatgggtc 1440
gggataatat ccgagtattt tcacagcgcg tttggaattg ttttgttcgt tctgctcgtc 1500
ctgaagcgca gcacatggac tctaatgatg gattct                           1536
```

<210> 87
<211> 426
<212> PRT
<213> Drosophila

<400> 87

```
Met Arg Leu Pro Trp Val Ile Phe Cys Thr Val Leu Leu Leu Ile Phe
 1               5                  10                  15
```

```
Thr Asn Asn Ser Asn Ala Asp Ile Pro Gly Cys Asn Tyr Tyr Asp Thr
            20                  25                  30
Val Asp Ile Ser Tyr Ile Glu Arg Gln Asn Asp Ser Tyr Leu Tyr Asp
        35                  40                  45
Asp Ile Glu Ile Pro Ala Ser Leu Thr Gly Tyr Tyr Glu Phe Arg Gln
    50                  55                  60
Phe Gly Asp Gly Ser Ile Thr Pro Ile Glu Lys His Leu Arg Ala Cys
65                  70                  75                  80
Val Cys Ser Val Arg Pro Cys Ile Arg Ile Cys Cys Pro Ala Lys Asn
                85                  90                  95
Phe Leu Ala Asn Gly Lys Cys Asp Asp Gly Leu Lys Glu Glu Leu Ala
            100                 105                 110
Arg Phe Lys Pro Tyr Ile Tyr Phe Thr Tyr Met Asp Leu Gln Ala Arg
        115                 120                 125
Val Pro Leu Thr Asp Met Ala Ile Ile Arg Asp Glu Phe Phe Asp Cys
    130                 135                 140
Asp Glu Met Ile Tyr Ile Ser Asp Phe Asn Tyr Phe Leu Glu Glu Val
145                 150                 155                 160
Ser Ile Gln Ile Phe Asn Lys Cys Gly Leu Ile Val Trp Phe Gln Asp
                165                 170                 175
Gly Lys Phe Trp Val Thr Val Asp Leu Phe Met Glu Lys Gln Asp Tyr
            180                 185                 190
Cys Leu Tyr Arg His Asn Phe Asp Ser Asp Phe Pro Lys Ser Met Trp
        195                 200                 205
Ile Ile Arg His Arg Cys Thr Ser His Ile Ser Pro Gly Ser Leu Glu
    210                 215                 220
Ile Leu Ile Ile Thr Met Ile Cys Phe Val Leu Thr Ile Ala Val Tyr
225                 230                 235                 240
Leu Tyr Ile Lys Lys Leu Arg Asn Val Thr Gly Lys Cys Ile Val Cys
                245                 250                 255
Cys Ile Val Ser Arg Phe Ile Gln Cys Leu Ile Met Ile Leu Asp His
            260                 265                 270
Leu Asn Leu Leu Asn Gly Ile Cys Ser Pro Ala Gly Tyr Ser Ser His
        275                 280                 285
Phe Phe Arg Met Ala Ser Asn Leu Trp Leu Ser Val Ile Ser Tyr His
    290                 295                 300
Thr Trp Lys Val Leu Thr Ser Leu Asn Arg Val Asp Pro Asn Tyr Arg
305                 310                 315                 320
Phe Leu Arg Tyr Asn Ala Phe Val Trp Ser Thr Ala Ala Ile Met Thr
                325                 330                 335
Gly Ser Ile Tyr Ile Val Asn Gln Ile Trp Glu Asn Asp Pro Ser Lys
            340                 345                 350
Trp Asn Trp Leu Pro Leu Val Gly Phe Ile Arg Cys Ser Val Lys Asp
        355                 360                 365
Trp His Pro Ser Val Trp Ile Tyr Ile Ser Gly Pro Ser Leu Ala Leu
    370                 375                 380
Ser Thr Phe Asn Val Ala Met Phe Ala Leu Thr Ala Ile Tyr Ile Arg
385                 390                 395                 400
Lys Val Lys Gly Gly Ile Asn Lys Phe Thr Asn Glu Glu Glu Gly Arg
                405                 410                 415
Ile Asn Cys Ile Asn Phe Asp Ser Gln Thr
            420                 425
```

<210> 88
<211> 6238
<212> DNA
<213> Drosophila

<400> 88

```
tgaacctgca ccaccagcca agaccccagc caagactccg accaaggcag ctgctccagc 60
cgctgttgcc agtcccgcag ctgcggcaac gtcggcagac gcctctccat ctccggccaa 120
gaaggcaacg cctgctcgcg ctgccgccgg agccaaggcc acgccgcagc gacaacgcgc 180
ccgcggtcgc tacaatcgct actaagtaga cgtcgaagga attcgtaatt taacataaca 240
taaaccttaa tatacaagaa aaaatttaca ttcacgccca ggaggaggat gaagaaagaa 300
cagatcaaat cgatctcttc ccgcaatttg acatttgcat cccaatttgc atatacactg 360
aatgaatgga aactgagaag aagacggtca acttttagtt gaaacagttg aatatattgt 420
acgacgaagc ttcttgaaac tgacgaggga gaacacaaac aagttaattg aatcgagaat 480
cgtattacct acccgattcc cccctcacag aacgcatacc atgccagcca gcgaaccaac 540
cattcaactt atgtaatcat ataaattcta taataaacaa agaaacaaaa ctagttgtaa 600
aacaaacacg attcctgtgt gtcattgcgg gatatggaaa ttggctttca atctgttgca 660
aaatgagatt cagtagcagt gtatatatct atgagggctg tacgatttcg ttttagcgca 720
caattaatac atttaatact tgatcttttt acacaagcgt tatatttaat gctacaaatc 780
ggtacagaga ttaagtgagg tgagaataca tacatataag gaaagccttg atatagtaag 840
atacgaagtg ggatatattt attatttaca gcttacgaga ctaggcgcta ctcccaggaa 900
tgctccgagg gaggcggaaa cgcttttttg agtgagatcc cactgagtcg ggaaatggga 960
ttaagtactt gggggcgcgc ggtatgaaga tcgagtggaa tcagttgaat tccagagtat 1020
tcgcctccgt catcgtcgta cgcatgcggc tctggtcgct ctgggaacga cgaagtccgc 1080
ccggacgcca tccgctcctc cagcgcttaa tgcatcccag caaggaccag ttgatggact 1140
tggagcgact cagcacatag acatacggat ccgaggtgaa gtgcagcgcc gtaagaacat 1200
cggcgatgat gaagaacttg ttcgatgcgg gcacccgatt tggagctata gccaacggga 1260
tggcgatcta aaacggagat aattgttgat aaaggccatg taaatttatt atattatatt 1320
agatttgcat aatacggaat ggttgaagat actagacgtt tagtagttac tatctgcaaa 1380
tttagaatgg tatccaccca gtactcacca tctgtggcat ccagcagatg acgaacgaga 1440
tgctgaggaa cgccattagt ttggcgaact tgatctcgac tggcgaggag tcggtggccg 1500
ccattgtgac gctaacgctg tgcctgtact gccttgccgg ctgtacactg cgatggctgt 1560
ttccgcttcc ggtgctcagc tgcgtctggt agagggtggt gccgctgctg ctctcgggat 1620
cgatgctgat ggcactgttt ttgtccctgg aaaccaggtc gtagtgcata tgccgcttgg 1680
ccgtgcgact ccttccgatc acacagagca aggtgtgggc cacgaagagg ttgcaggcca 1740
cgattacaat gcacagcagg gtacctggaa tacatcgatg gtcaaaccaa tatttgaaat 1800
gtaaccaatt tgataggaac ttaccgaaga ccataaagag caccgcatac gtcttgttcc 1860
acacgcccgg cgcatcgcga tatcgtatac acttaagctg gtccggattg gattcatcga 1920
tgtatgcacc aaaaccaacg aatggcagga acgtgatcac cacggcgatc atcagaatgc 1980
tgttgatgct cttgcgaatg agctcgtagg taatgtgcta tggcgagaga gatttgtgaa 2040
tatttaatga atgagttggc aatgtttgtg atttcctgtg tgcacttgta tcatttctta 2100
aaattaaatt acactaagaa taaaatgcat agaaaatgtt atttattata catatattgc 2160
aatcaagttc ataccttgtg gtagataaag ggtctggcta gggccatcca tctttctgcc 2220
gccataactg ctgcaatgca tccagaactg agcccgaaaa atcgccaaac cacgagtccc 2280
acgcaatcca ctcggatgaa cgactgcagg acttcctttg agagatacat tttcagcagc 2340
gtcgtggtca gcatgcccag tagagccacc agattgttgg tcgccaggca tctgaaatag 2400
ttgagagctc cgtaaatttc ggatagctta gctttatggc atttcgaagc tcataaatcg 2460
gaaagtaaat gcttttcctt gccattagcg cccaaattgc ctacagttcg ttgggttctt 2520
taataagcta cgaatataag ttgggtataa gatcgataaa atgtaataac gaattgcttt 2580
taattaattg ccacttttat aaaccgaaac tttcaatcaa gatacctcaa agctgctcca 2640
tatggttgta caactatcga ttttggttat ctaagacgtg actcggtcct taaacgcttt 2700
gattcgaacg ccgaccaata gcaatgaata attattttgg ttcgcaaaat gctcaattat 2760
aaaaaggaaa tatctattat atggatattt aaaaatttaa ttcatttttt tgaatgttta 2820
tattataaag gatgctcctc aaatttctcc attgaatatt gtgtatattc actttatatt 2880
cgttcatata tttaacattt gtatatttat actaatataa actaaagtac agatccctgt 2940
atttttatct acttagtgaa atatgtgact gtcaatggtg acttagatta agtgtttcct 3000
tttattggta tcataaaata ttgcatttgt ggggcacgaa tgtgtttcga tttgattgtt 3060
tttggggcgc cgttgacaca ggaaactcat ttgtctccgt gtgaaatgaa actatatgaa 3120
tgtagggctt ttttccaatg gaaatatatg tatatatata taaaacacac gcttggaata 3180
gaacacaaaa atgggattga caattcccgg ggattggcct aacagtccgt gcatcgttag 3240
gcatcgcaat gactttattc atttttgttt tcattatttg tttgtgcaca ttgcaccttt 3300
atacctgatt tcagcgaaca agcgattaaa atgcatttta tggtgcattg taaaagtttc 3360
gccaaaagct aaaatgtcgg aactgttcat tacagggcta aaatgtctaa atcttcaatt 3420
cgacgatttg cagtgcgaaa tgcaaaatcg ttttgctaac ttgtaaaggc aaaatcattg 3480
tttactctag caacgggaaa atacacacag ctattccgct tttcgttcat agtcaaaaag 3540
catgaaggga taccctgatt tatcgacacc gctcaagtgc cattcaccag cttcattcaa 3600
```

tcaatcaaat tacagatcaa aggctctgca tgaaggaggc ctgatgcctg aagtggaatt 3660
cgcgttaatt tatgcttcct tgacggctca aattgctggc gacagcaatt tacgcatatt 3720
acgcatacgc actgtcgtct cagcaaaaaa aaaatgttgg gaaaggagca aatcgcaaac 3780
aagctcactg aattcgccat gggccaatgc tgataaatgt aatataaatt gtgacaaata 3840
tttggcttat gaactaaaag aatttccgtg tgtgtcatag attttatcac ttctttacag 3900
tacacaccct accgaatgat taacacgatg tttaaaatat caacaaagtc tattatgagt 3960
attcgtattg agaagcgaat attttacaag aaattattat tattatatta atattaaaag 4020
gaatttggtg aaatacacaa gttttatttc actcatattt acaaaggaat agattatcat 4080
ttttgagtat tattatatct accagttttt tgcacctaag caaagttaat ttaaatacat 4140
attcatataa atcttttccc cagtatgcac atttgccaat caagtcaacc accgatcgcc 4200
gtcttcagtt acaccataaa gcttttaatg cattcgcaat tggctgcatt cccacgtcgc 4260
cgataagatt gacagaaaac agactggaat atggaatatg gacaaaatct caatgcccaa 4320
aacacatggc caacgaacta attccgtcat ttgataagat atgcaaaaat aaaacagcaa 4380
aaaccatgcc aagcaaaaca attcaaacta tgcaaatcag acccgcatag acatattata 4440
gagcaccaaa gtggccgcag ttgagagcag gttgaataaa cttacatagc catgaactgg 4500
acacaaacag tggcaaacat taactgcatg aaatgagtag ctacacgctg agaaaatatg 4560
caaaatgtca gctcaataat attattaaca ataatactat atggttttgc atgcttaata 4620
tttatgcaaa tcaaatagaa tttggtgaac tacacaacta taaattacta ctaagatggg 4680
tacattttga aagattttag ggcaagcaaa acctactttt tgcattaata tattattatt 4740
tccactatgc atagctgcac ttaccgcagc atgagcgtgt atttgctgtt tttgttgagc 4800
ttcttgcgcg ccaaaatgaa aagggccagc gaattgccaa aaacgcccag aaccatgata 4860
ataacaccaa taatcaggcg atttcgattt gcatagagcg ctttcggagg cgccacagta 4920
gttgagttgt gcggcatcaa caggtcgagc acaggtgttg tggtttccat tttgccggta 4980
tgataattgc gttcgattgc aggtgattca ggtgaaaatc taactggccg caaacgatat 5040
ttctttttca taaatgagcc tgccgcaaag gcaaggcagc tcgcaaatgc actggtagat 5100
ggaaacccgt ttcgatagcc aaattgggac tttttcccta agacacttaa ttaatgatta 5160
gcgcacgcga cacttgatca accaacgacc gaccaattcg aaggcgattc gattcgatcc 5220
gaggagcgcg tccgcagtca gccaggcgtc gatttagaat gactgccgac cgtgccgagc 5280
agagcactcg tcggctgctg cccatagctc caagctatcc aagccatcat atacacgtgg 5340
gggtggtggt aggggccagt tcggagcgag gccagttcgg ctggcgccgc cgccggtcgg 5400
atcgcatcgg atgtagcata cagcacatag acaccccccaa gcgaagcggc cacccacatc 5460
cacgcccagc gatggcgacc accccatcga aaattcacag ctcgtgccag cgcctttgcc 5520
ttttggacgt gtgcgaacgg tcggaatagc acgacgaatt ttattgtttt atactctctc 5580
cacctgcggc tcactttaat ttcactttca actggatgta atattaatta tgcagttttt 5640
aagcttccat ttgattttat acagttgact aaattagaat aaaataatat actcaaaagt 5700
ttacttcagc aatatattta ctcatgtttc tttgtcattg aaaaattaac aaaacaaaaa 5760
aacagctatt tcatagatag attaattatt attattaaga gtcttcagat tcgtcatctt 5820
tagtatatgg tgtgtcattt gaaaacctat ttttgaggaa tatatgaact gaaatctgcg 5880
tatactttaa ttactgtaat gtattaaagt attgtttctt ttatacaaaa agtctaaaaa 5940
tataagcaaa caaattattt tactgaaaat ttgtttttta tgccactttt aagaaatttt 6000
gcgaattaaa catatatttt taaagctagg agcctggtaa ttatagtagt ttagctttaa 6060
ctctttgctg aaaatatttc caactaattt atttgttttc aacagaatgc attacgaaag 6120
ataaaaaatta aatttgcagt tactaagcgt tgcagttgca ccttttctta ttcttatctt 6180
cttagtatat agtactgttt tgttgggcaa aaatgtttat tgtttagttt caactgcc 6238

<210> 89
<211> 1507
<212> DNA
<213> Drosophila

<400> 89
actgcggacg cgctcctcgg atcgaatcga atcgccttcg aattggtcgg tcgttggttg 60
atcaagtgtc gcgtgcgcta atcattaatt aagtgtctta ggaaaaaagt cccaatttgg 120
ctatcgaaac gggtttccat ctaccagtgc atttgcgagc tgccttgcct ttgcggcagg 180
ctcatttatg aaaaagaaat atcgtttgcg gccagttaga ttttcacctg aatcacctgc 240
aatcgaacgc aattatcata ccggcaaaat ggaaaccaca acacctgtgc tcgacctgtt 300
gatgccgcac aactcaacta ctgtggcgcc tccgaaagcg ctctatgcaa atcgaaatcg 360
cctgattatt ggtgttatta tcatggttct gggcgttttt ggcaattcgc tggccctttt 420

SUBSTITUTE SHEET (RULE 26)

```
cattttggcg cgcaagaagc tcaacaaaaa cagcaaatac acgctcatgc tgcgatgcct 480
ggcgaccaac aatctggtgg ctctactggg catgctgacc acgacgctgc tgaaaatgta 540
tctctcaaag gaagtcctgc agtcgttcat ccgagtggat tgcgtgggac tcgtggtttg 600
gcgatttttc gggctcagtt ctggatgcat tgcagcagtt atggcggcag aaagatggat 660
ggccctagcc agaccctttta tctaccacaa gcacattacc tacgagctca ttcgcaagag 720
catcaacagc attctgatga tcgccgtggt gatcacgttc ctgccattcg ttggttttgg 780
tgcatacatc gatgaatcca atccggacca gcttaagtgt atacgatatc gcgatgcgcc 840
gggcgtgtgg aacaagacgt atgcggtgct ctttatggtc ttcggtaccc tgctgtgcat 900
tgtaatcgtg gcctgcaacc tcttcgtggc ccacaccttg ctctgtgtga tcggaaggag 960
tcgcacggcc aagcggcata tgcactacga cctggtttcc agggacaaaa acagtgccat 1020
cagcatcgat cccgagagca gcagcggcac caccctctac cagacgcagc tgagcaccgg 1080
aagcggaaac agccatcgca gtgtacagcc ggcaaggcag tacaggcaca gcgttagcgt 1140
cacaatggcg gccaccgact cctcgccagt cgagatcaag ttcgccaaac taatggcgtt 1200
cctcagcatc tcgttcgtca tctgctggat gccacagatg atcgccatcc cgttggctat 1260
agctccaaat cgggtgcccg catcgaacaa gttcttcatc atcgccgatg ttcttacggc 1320
gctgcacttc acctcggatc cgtatgtcta tgtgctgagt cgctccaagt ccatcaactg 1380
gtccttgctg ggatgcatta agcgctggag gagcggatgg cgtccgggcg gacttcgtcg 1440
ttcccagagc gaccagagcc gcatgcgtac gacgatgacg gaggcgaata ctctggaatt 1500
caactga                                                           1507
```

<210> 90
<211> 412
<212> PRT
<213> Drosophila

<400> 90

```
Met Glu Thr Thr Thr Pro Val Leu Asp Leu Leu Met Pro His Asn Ser
 1               5                  10                  15
Thr Thr Val Ala Pro Pro Lys Ala Leu Tyr Ala Asn Arg Asn Arg Leu
            20                  25                  30
Ile Ile Gly Val Ile Ile Met Val Leu Gly Val Phe Gly Asn Ser Leu
        35                  40                  45
Ala Leu Phe Ile Leu Ala Arg Lys Lys Leu Asn Lys Asn Ser Lys Tyr
    50                  55                  60
Thr Leu Met Leu Arg Cys Leu Ala Thr Asn Asn Leu Val Ala Leu Leu
65                  70                  75                  80
Gly Met Leu Thr Thr Thr Leu Leu Lys Met Tyr Leu Ser Lys Glu Val
                85                  90                  95
Leu Gln Ser Phe Ile Arg Val Asp Cys Val Gly Leu Val Val Trp Arg
            100                 105                 110
Phe Phe Gly Leu Ser Ser Gly Cys Ile Ala Ala Val Met Ala Ala Glu
        115                 120                 125
Arg Trp Met Ala Leu Ala Arg Pro Phe Ile Tyr His Lys His Ile Thr
    130                 135                 140
Tyr Glu Leu Ile Arg Lys Ser Ile Asn Ser Ile Leu Met Ile Ala Val
145                 150                 155                 160
Val Ile Thr Phe Leu Pro Phe Val Gly Phe Gly Ala Tyr Ile Asp Glu
                165                 170                 175
Ser Asn Pro Asp Gln Leu Lys Cys Ile Arg Tyr Arg Asp Ala Pro Gly
            180                 185                 190
Val Trp Asn Lys Thr Tyr Ala Val Leu Phe Met Val Phe Gly Thr Leu
        195                 200                 205
Leu Cys Ile Val Ile Val Ala Cys Asn Leu Phe Val Ala His Thr Leu
    210                 215                 220
Leu Cys Val Ile Gly Arg Ser Arg Thr Ala Lys Arg His Met His Tyr
225                 230                 235                 240
Asp Leu Val Ser Arg Asp Lys Asn Ser Ala Ile Ser Ile Asp Pro Glu
                245                 250                 255
Ser Ser Ser Gly Thr Thr Leu Tyr Gln Thr Gln Leu Ser Thr Gly Ser
            260                 265                 270
```

Gly Asn Ser His Arg Ser Val Gln Pro Ala Arg Gln Tyr Arg His Ser
        275   280   285
Val Ser Val Thr Met Ala Ala Thr Asp Ser Ser Pro Val Glu Ile Lys
    290   295   300
Phe Ala Lys Leu Met Ala Phe Leu Ser Ile Ser Phe Val Ile Cys Trp
305   310   315   320
Met Pro Gln Met Ile Ala Ile Pro Leu Ala Ile Ala Pro Asn Arg Val
                325   330   335
Pro Ala Ser Asn Lys Phe Phe Ile Ile Ala Asp Val Leu Thr Ala Leu
            340   345   350
His Phe Thr Ser Asp Pro Tyr Val Tyr Val Leu Ser Arg Ser Lys Ser
        355   360   365
Ile Asn Trp Ser Leu Leu Gly Cys Ile Lys Arg Trp Arg Ser Gly Trp
    370   375   380
Arg Pro Gly Gly Leu Arg Arg Ser Gln Ser Asp Gln Ser Arg Met Arg
385   390   395   400
Thr Thr Met Thr Glu Ala Asn Thr Leu Glu Phe Asn
                405   410

<210> 91
<211> 5340
<212> DNA
<213> Drosophila

<400> 91
tacactcatg aatattcaaa caaaactctt caacttgtct ggcatatatc catagacaat 60
tacgcgggta tgtatgctgg gttttgtgtt cttaaaatcg tatatctgct gcaggattga 120
catcgaactg cttaacccca cgacggcaac cattcgtcga aaaagtttga acacagttta 180
tatttttgca atattctaag tactataacc taccaaaata actgtaaaag aaagtttcta 240
tgaacgtttg aaatcaaaat taggtccgat tttgcgtagt tttcgacaaa gtgtcaccgt 300
caaaagggct tcgactggca tatccctcaa agggttaaca tgcttttctt tggtcttctt 360
atggcctacc cccttcatat gtctgctttt cattgatata ttctgcggaa caacaaaaga 420
agtattattc tatattggga tatgctttag agttttggtt tgtagggttt tctagtctcg 480
gtaaatgtgt atatcttatg cttaatttat tttattagct atcatatcat atactttgca 540
tggttatcat agattggtta ttaacattaa aatatttatt gtctaaaaca tttcgccccg 600
cttagtgcca ctctagacct atgtatgcat cccaaatcgg attgcttgtt ttcccccccca 660
acgattctaa ctatagcaca cacctttggc tacttacaga taactcttcc ctgtccaccg 720
attaaatttt gaatttaata agtgtttgat tctaatactt tgctgatggt cgaaatccag 780
cagatattcc attaagcgga gcactttcgt tgggtatgta tagaaaactt atgggcctaa 840
ggaggtgtaa ggggtttggg aatggaaaag tttgcaaggc gccagactct cggggctctg 900
aagtgtctgg tcctttgcgc aagtgcagtg caaaatggaa atctatcagg gatttaactc 960
aaccaccagc agcttaacca cttcagatcc ccaaaccatc ttaggtcatg tgcagatccc 1020
ccttgaggat ccgcgtaaag gtcttcttga actgctgatt ggccagagcg tagcaaaaag 1080
gattcatcgg gctgttggcg tagcagagga agtacgagaa catgtagagg tgttcattga 1140
tgcagggcgg atgccggcag aatccctcga cgagggccaa cacatggtag ggtgtccagc 1200
aggccacgaa gcagcccaaa atgaacgaga tcgtgcgaaa cgctttgcga gcacgattct 1260
ccgacttgga cttctgccgt ccacccactc caaggatcag ggtagagcc ttcttgctct 1320
tgaagtgctt tcctatcgaa tgaatgaatg ccctttttgct ggagtttcgg ctggccgcca 1380
aactcttctc ctccctgttg gcatcgtcat cccccgcgtt ctttttggtg gcaccaacac 1440
cgtcttgttc cccgcctcca cccacaccgc ccgtactgct ggtggtgccc gaggtggagg 1500
gcatcttgtt ggtattactg tccgaggttt ctcgctcctc aatcctttcc agtggcaccg 1560
gtggtgttac acccgccgcc gactggggcg ttcccgtttg tggcgttgaa gtctttacct 1620
tgggtgagag tggcattggt ggcatcgggg aatagagtat ggcatccttg ctgctggtct 1680
ccataggact ggaggcccgg cgatggggac agtccgcagc tagctttccg ggcacatcca 1740
acaggcagat ggcctcaatg cagctctccc gcctcacctt agcaaagtcc tcgattcctg 1800
tgtacgagac gagaagacgt ttttcaatgg ccaccgcctc ctcgatactc tgactctgac 1860
tctggttgtg attcggattc aggttgaggc tcaggctttg gctgggactc tggttttggc 1920
tgggagtttg actgaggctc tggttcaggt tcgattttgg tggatttcgg cgtgcgggcg 1980
gtggtggagg caggactggt ggtttcccct ggacgctgtc attcggggag gtcgtcgagt 2040

tggccgtcac actggccagc atggcggagc tctcgtccat atagcgtagc tccgagttcg 2100
tcatcaggtc ataggtcagg gccagttcac tgttaccgga actattccca aagggacgat 2160
ccgagtagtc atcggagagc ggggtggggc actgaaaggc gtcgggaggt ggcaaaatgg 2220
tctgcaccag accctgatgc acttgatcgc tagggacctc gatgggtgcc aagtccacgg 2280
gggacagggg atcactggtg ttgagtaatg ttggaggagc ggacaggtcg accaaacgat 2340
catgctcggc tatctgggac agtgtggagc aggatgtgcg cttgggcctg tccagcggtg 2400
ggttcggctg gttttggtgg tgctggtggt gcggttgttg ctggttctgg tggtggtgca 2460
gggagcgcat ctcttccttg ctctgcgccc gctgtaacgg actctggttg tgatggtgct 2520
gttggtggtg gtggtactgg tgcatggcct cgatcgactt ggagttagtg gtgctcttct 2580
gcaagttacc tttgccacga gtggccagca gggcggcctg tgccccaaag accagaccaa 2640
tcgatgaccg cttcctcatg ttgttcaatt tctgctgggt gatcaactgc tgggcctggt 2700
tgacggatga ttcctcatcg gaatcgaagg caggactgct ggatctctcg ctcttttccg 2760
catccacggc tgcctcacgt tttttcgcct cctcctggat ttttgctgcc gacgctctcc 2820
tctgttccgg tgtcatcttc tccacttcct cggtggccac ggccagagga tttgcctgac 2880
cggaaccact taatctctta atcactgtgg tgcaggcact atccaaatcg gtttgatctc 2940
ccgcaagttc cactttggtc tttaggattt tctcctcaat cacaccaatg ccagctgcat 3000
gacctgccat tccggacatc gctccagcac tcagtgccac catggattgc attttccgtt 3060
gcttcgcctc actgcgcttt tgcatatcat aggctgtttt gtagatcccg gcatataaga 3120
caaaaagcac aattagtgtg gtccaatagt agccaatgat tagggccgta ttaaagatcg 3180
gatccttgag aaactgaacg gcgcactggc caggaagtag atcccgcttg ccggtgaagt 3240
gctcccatcc aaaaatcgag atgaagaaga gcagagcggg aatgatccaa gtgattgtga 3300
ccatatatat cactctggtt ctcgtcctcc aactccggta cttggcggca atcttcacag 3360
agcagtagcg atctatcgtg atcaacagca ccgtgtactg ggacaccaga cacacagtat 3420
agtcgacgga aagccagaga tcgcagagca tcggacccag gtcccaataa cccttgagca 3480
cgtatatcgt gtagaagggc atcgaaacgg ttcctgaaat atgggtttat agtacaatta 3540
ataatttttt ttttttgaat ttaaggtaaa atagttacta cctattagca tatccgtggc 3600
tgctagggag gcaataaagt agttgctggg ctgccgaata ttcctgtcca cgatgaaggc 3660
cagtagcacc aggatattac cgcccaccgt aagaatgata cagatcgcca ggcatatggc 3720
tatcaggatc gtctgccaca gtgcaaatgg tggcaaaaca ggaccaagaa tctccgagtc 3780
ggaactggat cccgacatgg tggaaccgcc atcagaggat tgatttctaa cagccaacag 3840
aaggttcgta ttgttcacat cgaagttcct gtaccattgc agctgttcat taaccggcca 3900
ggaattgaag tcactgctct cgccatagat cttggccgcc tccaacagac gtttgcactc 3960
cgtgtgcgtc aggactcgct tgaaaatgaa gctgcacagc ttatcccgct cccttctaag 4020
gctcctggaa gtgcgaaagt aaatatatgc ttaattagag ttgacataaa actatctcta 4080
tctctatctc tatctctatc tctatctcta tctctaaact aatgtccttg aaaacaaatt 4140
ctaacaatgt atctttttaa gatctccatt ttctaaaagc tttgccactg aaattactat 4200
tctaataggc gcctttattc tattaagaca actcaccagt cggtgatcca cttgttcttg 4260
atggtataga tggagtcgct ggagttgcca ttcaggccag tcagctcctc ctcgcccgcc 4320
gacttgttaa gcgaccgcat ccttaaccgc tccgccctct tcccgctctg gccacgtccc 4380
gctccctgct ccccgcactt cgcctctgcg ctccacaacc tctgcggttt ttatggtaat 4440
ttcaagttgt tcctccttaa tgtgctgtcg acggcggttg tctcggaatg cttacattgc 4500
ctccatggaa tgccaagact tgtctccggt caggtgcagc atctgcaaga ggcaacaaaa 4560
gcagccacag tcttggttag cacttgaaaa aatatagagg ttattaatat tatatgatat 4620
aatatatata tgcatataaa atatgatata atatatatat gcatataaaa tatgatataa 4680
tttatatatg tatacataat atgataaaaa gtctgaattt cagaattctg ccagaagtcg 4740
ttttccatcc aaaaattaat tgttttttaa tcgaattttt ggtaataata atccgaataa 4800
ggaatggaat atttcgttga gatcgtaaat aaattttaaa tccgtgtgtg tttgatttgg 4860
aaccagtcgt tttccatcca aaaattaatt gtttttttaat cgaatttttg gtaataataa 4920
tccgaataag gaatggaata tctcgttgag atcgtaaata aattttaaaa tatatgtgtg 4980
tttgatttgg aaccatttat ttttttcacg atgttgccac tttccaaaaa ttcgaaaatt 5040
gacaaaaaaa atatttttcc gttttagttg cgcgaccatt tttagccaaa taatagtaaa 5100
aaatctattt gaaaaataca ttgtcttcat tagatggtgg aaaataatga aaatatgaaa 5160
ataatggaaa ttaaattaaa tcaaaattag taatgaaaat caagtattta ctaaataaag 5220
aaactttctt ttcgtgtagt cttcgtggac gtcttcgctt ccgccaagtt taaactaatt 5280
ttgcaaccgt taactggcaa ccggcgactg gcgatgcaat ttgcttgtgg cagttgcttc 5340

<210> 92

<211> 3060
<212> DNA
<213> Drosophila

<400> 92
atgcggtcgc ttaacaagtc ggcgggcgag gaggagctga ctggcctgaa tggcaactcc 60
agcgactcca tctataccat caagaacaag tggatcaccg actggagcct tagaagggag 120
cgggataagc tgtgcagctt cattttcaag cgagtcctga cgcacacgga gtgcaaacgt 180
ctgttggagg cggccaagat ctatggcgag agcagtgact tcaattcctg gccggttaat 240
gaacagctgc aatggtacag gaacttcgat gtgaacaata cgaaccttct gttggctgtt 300
agaaatcaat cctctgatgg cggttccacc atgtcgggat ccagttccga ctcggagatt 360
cttggtcctg ttttgccacc atttgcactg tggcagacga tcctgatagc catatgcctg 420
gcgatctgta tcattcttac ggtgggcggt aatatcctgg tgctactggc cttcatcgtg 480
gacaggaata ttcggcagcc cagcaactac tttattgcct ccctagcagc cacggatatg 540
ctaataggaa ccgtttcgat gcccttctac acgatatacg tgctcaaggg ttattgggac 600
ctgggtccga tgctctgcga tctctggctt tccgtcgact atactgtgtg tctggtgtcc 660
cagtacacgg tgctgttgat cacgatagat cgctactgct ctgtgaagat tgccgccaag 720
taccggagtt ggaggacgag aaccagagtg atatatatgg tcacaatcac ttggatcatt 780
cccgctctgc tcttcttcat ctcgattttt ggatgggagc acttcaccgg caagcgggat 840
ctacttcctg gccagtgcgc cgttcagttt ctcaaggatc cgatctttaa tacggcccta 900
atcattggct actattggac cacactaatt gtgctttttg tcttatatgc cgggatctac 960
aaaacagcct atgatatgca aaagcgcagt gaggcgaagc aacggaaaat gcaatccatg 1020
gtggcactga gtgctggagc gatgtccgga atggcaggtc atgcagctgg cattggtgtg 1080
attgaggaga aaatcctaaa gaccaaagtg gaacttgcgg gagatcaaac cgatttggat 1140
agtgcctgca ccacagtgat taagagatta agtggttccg gtcaggcaaa tcctctggcc 1200
gtggccaccg aggaagtgga gaagatgaca ccggaacaga ggagagcgtc ggcagcaaaa 1260
atccaggagg aggcgaaaaa acgtgaggca gccgtggatg cggaaaagag cgagagatcc 1320
agcagtcctg ccttcgattc cgatgaggaa tcatccgtca accaggccca gcagttgatc 1380
acccagcaga aattgaacaa catgaggaag cggtcatcga ttggtctggt ctttggggca 1440
caggccgccc tgctggccac tcgtggcaaa ggtaacttgc agaagagcac cactaactcc 1500
aagtcgatcg aggccatgca ccagtaccac caccaccaac agcaccatca caaccagagt 1560
ccgttacagc gggcgcagag caaggaagag atgcgctccc tgcaccacca ccagaaccag 1620
caacaaccgc accaccagca ccaccaaaac cagccgaacc caccgctgga caggcccaag 1680
cgcacatcct gctccacact gtcccagata gccgagcatg atcgtttggt cgacctgtcc 1740
gctcctccaa cattactcaa caccagtgat cccctgtccc ccgtggactt ggcacccatc 1800
gaggtcccta gcgatcaagt gcatcagggt ctggtgcaga ccattttgcc acctccccgac 1860
gcctttcagt gccccacccc gctctccgat gactactcgg atcgtccctt tgggaatagt 1920
tccggtaaca gtgaactggc cctgacctat gacctgatga cgaactcgga gctacgctat 1980
atggacgaga gctccgccat gctggccagt gtgacggcca actcgacgac ctccccgaat 2040
gacagcgtcc aggggaaacc accagtcctg cctccaccac cgcccgcacg ccgaaatcca 2100
ccaaaatcga acctgaacca gagcctcagt caaactccca gccaaaacca gagtcccagc 2160
caaagcctga gcctcaacct gaatccgaat cacaaccaga gtcagagtca gagtatcgag 2220
gaggcggtgg ccattgaaaa acgtcttctc gtctcgtaca caggaatcga ggactttgct 2280
aaggtgaggc gggagagctg cattgaggcc atctgcctgt tggatgtgcc cggaaagcta 2340
gctgcggact gtccccatcg ccgggcctcc agtcctatgg agaccagcag caaggatgcc 2400
atactctatt ccccgatgcc accaatgcca ctctcaccca aggtaaagac ttcaacgcca 2460
caaacgggaa cgccccagtc ggcggcgggt gtaacaccac cggtgccact ggaaaggatt 2520
gaggagcgag aaacctcgga cagtaatacc aacaagatgc cctccacctc gggcaccacc 2580
agcagtacgg gcggtgtggg tggaggcggg gaacaagacg gtgttggtgc caccaaaaag 2640
aacgcggggg atgacgatgc caacagggag gagaagagtt tggcggccag ccgaaactcc 2700
agcaaaaggg cattcattca ttcgatagga aagcacttca agagcaagaa ggctctaccc 2760
ctgatccttg gagtgggtgg acggcagaag tccaagtcgg agaatcgtgc tcgcaaagcg 2820
tttcgcacga tctcgttcat tttgggctgc ttcgtggcct gctggacacc ctaccatgtg 2880
ttggccctcg tcgagggatt ctgccggcat ccgccctgca tcaatgaaca cctctacatg 2940
ttctcgtact tcctctgcta cgccaacagc ccgatgaatc ctttttgcta cgctctggcc 3000
aatcagcagt tcaagaagac ctttacgcgg atcctcaagg gggatctgca catgacctaa 3060

<210> 93
<211> 1019

<212> PRT
<213> Drosophila

<400> 93
Met Arg Ser Leu Asn Lys Ser Ala Gly Glu Glu Glu Leu Thr Gly Leu
1 5 10 15
Asn Gly Asn Ser Ser Asp Ser Ile Tyr Thr Ile Lys Asn Lys Trp Ile
20 25 30
Thr Asp Trp Ser Leu Arg Arg Glu Arg Asp Lys Leu Cys Ser Phe Ile
35 40 45
Phe Lys Arg Val Leu Thr His Thr Glu Cys Lys Arg Leu Leu Glu Ala
50 55 60
Ala Lys Ile Tyr Gly Glu Ser Ser Asp Phe Asn Ser Trp Pro Val Asn
65 70 75 80
Glu Gln Leu Gln Trp Tyr Arg Asn Phe Asp Val Asn Asn Thr Asn Leu
85 90 95
Leu Leu Ala Val Arg Asn Gln Ser Ser Asp Gly Gly Ser Thr Met Ser
100 105 110
Gly Ser Ser Ser Asp Ser Glu Ile Leu Gly Pro Val Leu Pro Pro Phe
115 120 125
Ala Leu Trp Gln Thr Ile Leu Ile Ala Ile Cys Leu Ala Ile Cys Ile
130 135 140
Ile Leu Thr Val Gly Gly Asn Ile Leu Val Leu Leu Ala Phe Ile Val
145 150 155 160
Asp Arg Asn Ile Arg Gln Pro Ser Asn Tyr Phe Ile Ala Ser Leu Ala
165 170 175
Ala Thr Asp Met Leu Ile Gly Thr Val Ser Met Pro Phe Tyr Thr Ile
180 185 190
Tyr Val Leu Lys Gly Tyr Trp Asp Leu Gly Pro Met Leu Cys Asp Leu
195 200 205
Trp Leu Ser Val Asp Tyr Thr Val Cys Leu Val Ser Gln Tyr Thr Val
210 215 220
Leu Leu Ile Thr Ile Asp Arg Tyr Cys Ser Val Lys Ile Ala Ala Lys
225 230 235 240
Tyr Arg Ser Trp Arg Thr Arg Thr Arg Val Ile Tyr Met Val Thr Ile
245 250 255
Thr Trp Ile Ile Pro Ala Leu Leu Phe Phe Ile Ser Ile Phe Gly Trp
260 265 270
Glu His Phe Thr Gly Lys Arg Asp Leu Leu Pro Gly Gln Cys Ala Val
275 280 285
Gln Phe Leu Lys Asp Pro Ile Phe Asn Thr Ala Leu Ile Ile Gly Tyr
290 295 300
Tyr Trp Thr Thr Leu Ile Val Leu Phe Val Leu Tyr Ala Gly Ile Tyr
305 310 315 320
Lys Thr Ala Tyr Asp Met Gln Lys Arg Ser Glu Ala Lys Gln Arg Lys
325 330 335
Met Gln Ser Met Val Ala Leu Ser Ala Gly Ala Met Ser Gly Met Ala
340 345 350
Gly His Ala Ala Gly Ile Gly Val Ile Glu Glu Lys Ile Leu Lys Thr
355 360 365
Lys Val Glu Leu Ala Gly Asp Gln Thr Asp Leu Asp Ser Ala Cys Thr
370 375 380
Thr Val Ile Lys Arg Leu Ser Gly Ser Gly Gln Ala Asn Pro Leu Ala
385 390 395 400
Val Ala Thr Glu Glu Val Glu Lys Met Thr Pro Glu Gln Arg Arg Ala
405 410 415
Ser Ala Ala Lys Ile Gln Glu Glu Ala Lys Lys Arg Glu Ala Ala Val
420 425 430
Asp Ala Glu Lys Ser Glu Arg Ser Ser Ser Pro Ala Phe Asp Ser Asp
435 440 445

```
Glu Glu Ser Ser Val Asn Gln Ala Gln Gln Leu Ile Thr Gln Gln Lys
    450                 455                 460
Leu Asn Asn Met Arg Lys Arg Ser Ser Ile Gly Leu Val Phe Gly Ala
465                 470                 475                 480
Gln Ala Ala Leu Leu Ala Thr Arg Gly Lys Gly Asn Leu Gln Lys Ser
                485                 490                 495
Thr Thr Asn Ser Lys Ser Ile Glu Ala Met His Gln Tyr His His His
            500                 505                 510
Gln Gln His His His Asn Gln Ser Pro Leu Gln Arg Ala Gln Ser Lys
        515                 520                 525
Glu Glu Met Arg Ser Leu His His His Gln Asn Gln Gln Gln Pro His
    530                 535                 540
His Gln His His Gln Asn Gln Pro Asn Pro Pro Leu Asp Arg Pro Lys
545                 550                 555                 560
Arg Thr Ser Cys Ser Thr Leu Ser Gln Ile Ala Glu His Asp Arg Leu
                565                 570                 575
Val Asp Leu Ser Ala Pro Pro Thr Leu Leu Asn Thr Ser Asp Pro Leu
            580                 585                 590
Ser Pro Val Asp Leu Ala Pro Ile Glu Val Pro Ser Asp Gln Val His
        595                 600                 605
Gln Gly Leu Val Gln Thr Ile Leu Pro Pro Pro Asp Ala Phe Gln Cys
    610                 615                 620
Pro Thr Pro Leu Ser Asp Asp Tyr Ser Asp Arg Pro Phe Gly Asn Ser
625                 630                 635                 640
Ser Gly Asn Ser Glu Leu Ala Leu Thr Tyr Asp Leu Met Thr Asn Ser
                645                 650                 655
Glu Leu Arg Tyr Met Asp Glu Ser Ser Ala Met Leu Ala Ser Val Thr
            660                 665                 670
Ala Asn Ser Thr Thr Ser Pro Asn Asp Ser Val Gln Gly Lys Pro Pro
        675                 680                 685
Val Leu Pro Pro Pro Pro Pro Ala Arg Arg Asn Pro Pro Lys Ser Asn
    690                 695                 700
Leu Asn Gln Ser Leu Ser Gln Thr Pro Ser Gln Asn Gln Ser Pro Ser
705                 710                 715                 720
Gln Ser Leu Ser Leu Asn Leu Asn Pro Asn His Asn Gln Ser Gln Ser
                725                 730                 735
Gln Ser Ile Glu Glu Ala Val Ala Ile Glu Lys Arg Leu Leu Val Ser
            740                 745                 750
Tyr Thr Gly Ile Glu Asp Phe Ala Lys Val Arg Arg Glu Ser Cys Ile
        755                 760                 765
Glu Ala Ile Cys Leu Leu Asp Val Pro Gly Lys Leu Ala Ala Asp Cys
    770                 775                 780
Pro His Arg Arg Ala Ser Ser Pro Met Glu Thr Ser Ser Lys Asp Ala
785                 790                 795                 800
Ile Leu Tyr Ser Pro Met Pro Pro Met Pro Leu Ser Pro Lys Val Lys
                805                 810                 815
Thr Ser Thr Pro Gln Thr Gly Thr Pro Gln Ser Ala Ala Gly Val Thr
            820                 825                 830
Pro Pro Val Pro Leu Glu Arg Ile Glu Glu Arg Glu Thr Ser Asp Ser
        835                 840                 845
Asn Thr Asn Lys Met Pro Ser Thr Ser Gly Thr Thr Ser Ser Thr Gly
    850                 855                 860
Gly Val Gly Gly Gly Gly Glu Gln Asp Gly Val Gly Ala Thr Lys Lys
865                 870                 875                 880
Asn Ala Gly Asp Asp Asp Ala Asn Arg Glu Glu Lys Ser Leu Ala Ala
                885                 890                 895
Ser Arg Asn Ser Ser Lys Arg Ala Phe Ile His Ser Ile Gly Lys His
            900                 905                 910
Phe Lys Ser Lys Lys Ala Leu Pro Leu Ile Leu Gly Val Gly Gly Arg
        915                 920                 925
```

```
Gln Lys Ser Lys Ser Glu Asn Arg Ala Arg Lys Ala Phe Arg Thr Ile
    930                 935                 940
Ser Phe Ile Leu Gly Cys Phe Val Ala Cys Trp Thr Pro Tyr His Val
945                 950                 955                 960
Leu Ala Leu Val Glu Gly Phe Cys Arg His Pro Pro Cys Ile Asn Glu
                965                 970                 975
His Leu Tyr Met Phe Ser Tyr Phe Leu Cys Tyr Ala Asn Ser Pro Met
            980                 985                 990
Asn Pro Phe Cys Tyr Ala Leu Ala Asn Gln Gln Phe Lys Lys Thr Phe
        995                1000                1005
Thr Arg Ile Leu Lys Gly Asp Leu His Met Thr
    1010                1015


&lt;210&gt; 94
&lt;211&gt; 4172
&lt;212&gt; DNA
&lt;213&gt; Drosophila

&lt;400&gt; 94
gacatacgca ggcggagaat tgatttgatt tctgttattg ttgcttatca ctatcgattt 60
cgattattca aagtcccttt ttgccctgcg gcgaacggcg attgccctca gtttctgagt 120
tcctgcaaat caaatgcaat tatgtcgttc aatcgtctga atgctgaacg gtaaatcgag 180
taaacaagag tgggcggagg tggtctgaat aatttatgac cgcccattaa ttattactca 240
tagagctgtc ctatgactta tcgactttaa gccaattgga gagggggggag taaggctggc 300
agggcaaccc gatcaaaacg aaaaatcaaa gtgtgggtat cattgtgtat ttcgtaaatt 360
agaaggccaa tttggtggca attcctgttt gtgatacaat tcctggccgc cattgtaccc 420
gtggtacttc taggaattca ttaaatgtat ccaatttgct gtcaaggtaa tttacttggt 480
gaacaattaa gttcaagtag ctatgtttag tcatttcaaa gttatgtctt tacaatttat 540
tttatataaa tactaagtgc tattcttaat taagcatctg cacttaaggt tttctctata 600
aggagcctat tcaaatagtt ttcattaatg ttttcattct ttccctttct cgattgcagc 660
tccaacttcc ggcgctttgg tggcgtaaaa tgtttgccac caatatccgt tcggttgccg 720
tagaaatcgt cctggtggcg ccgtgagctc tcgtcctttt cctttcgcga gtcctgaatc 780
cccgtccctc gctcatttcc ggtgcgccgt ttttacattt gcaattcgaa tcggattgaa 840
atcggaatcg gaatcagaat cggcgcgagg acatttccat tccttcccat tctcttttta 900
tttttttttt ttgcgatctg cgttgcatat gcgagggcta attagcatgc ggcatttcca 960
gcaatcagag tagagcgcac aataggaata ataacgcgga atggaagagg atgtgtatgc 1020
ctcgctaggt gcctacaacg acagcggtgg cgacgattgg agcagctcgg agcacctggt 1080
cctgtgggag gaggatgaga cgcagcgaac gactgctaat gccaccagtc ggcataatca 1140
actgcatgtg gccaggtgga atgccaccgg caatgcgacc atcagcgcga ccttcgagga 1200
cgtacccttc gatgcgaaca actactgggc cctattggcc ctcgtcctcg tcctgggaac 1260
agcggccggg aacatccttg tctgcctggc catcgcctgg gagcgccggc tgcagaatgt 1320
gaccaactac ttccttatgt cgctggccat caccgatctc atggtcgccg tgctggtcat 1380
gccgctgggc atccttacgc tggtgaaagg tgagtccttt gtcccattca aagtccggaa 1440
cagaggtatt ttttgcagcc accttttgtt gagggactca atggacactt gagtaagtca 1500
cttgtgcgcc aggattggct tgtgcggcta atcttagttg gaaaagttcc ttttccataa 1560
tttcatagtt tactctaacg aaaatgagtt ggtggaaaac ttgatgaaaa ttaatattaa 1620
ttacttgggg caagggaaaa aagagtgaga agttttgata gttgttaaat ttttaaggct 1680
tttaaaaagt gacgaagaag atgcaaggaa acgcaaagat gtagggaata tacaaaatta 1740
ttttaaatct taagatttct tatatttgca atttaaatgc ggtaaatgcg aatcaattga 1800
ataatgtaca gggaattgtg aaaaaatatg gatatttacg gattcctttt gctgcaactc 1860
aatccgaatg ataatctctc caagtggaag agtataataa tcaaagccag cgacagaaac 1920
gtggcaaaga cattcaggcc aggacctcct gcgatgtcgc caacacgagg cacgagttgc 1980
atttaatgaa gctgctggtg ggaatggcgt tggcggtgtg cagattgcat tatgtctcag 2040
atatttcgat caaagctggc aaacgtaagt aagcccggat agaaaaatgg aatcgctggt 2100
gtctatgggc cggcagggcg gagcggagct ccatcgttag acaatttaat aagcgcggcg 2160
cctttgtagc cagcctggcc aggacgaacc gctggcgcag acagggaagc ccagaaacaa 2220
tgccagcaag atggcaacgg aagcttccaa agcctttgca ggcgccggag caactgcagc 2280
ataagaattg cattaggggc ggcaaacgag cgctggtgaa tttcgctttg acttagttcg 2340
gatctggaaa catctgccga cacattgctg tcctgcaatc tgcactgaaa aaaaaaacaa 2400
```

aatatattta ttcaattatt tctatactta aagtaaaagt catcaagttt aaagctaatt 2460
gtataatttt tttctttcga ttttttcgat tgagcaataa aatatggttc atatatgtta 2520
atacttacat tttttttaag ctaagtaact tgaagctagc tacacaaata ttataacata 2580
ccagacagga aatggaaaag ctatttccgt gactttcaaa ggcgttgctt atttcgttgc 2640
tccatttatg tgagtgcttc ttccggtggc agttaaacgg tatttcttcg attttcgaaa 2700
acgaatggct gatttgagga ctgcatttat ctgagtgctg cttccgcctc ctgctgccgt 2760
cctttttgtg ccaacccctg cagatgttgc aacactgatg acagttttgt ttgtgcgctt 2820
tgggagccat tacaaacaca aacacaaaca caaacacgag cacgggcacc aatacaagca 2880
gaaagggcga cgcacacaca cgcaaacaca cacctggagc aatccttggc aaagtgggca 2940
atgtagcaag aaggaggacc aacaaatggc tgcttgttat atcaagtgtg aaaaatattg 3000
cagtggcagt gacaggagca gttgttgctg tagttgcacg ggcacacgcc cgccaactgg 3060
aggcaagaaa ggacgaaatc gggatttgcc tgctgtcaag ggctttgcag agggtcctgc 3120
gacagaagct gaaatctcca gcgaacagca ggcattagat cccaaggatt aagctcttac 3180
aattcaaagg cacaatggcg caagtagcaa ctggatatga gtaagccgtg cagtacatta 3240
attgcatcgg ttacgtaagt aacatgtggc atacaatgtg gcattgaaat ggcctatata 3300
ttctattctc tgttgataag tatgcaaatt caactttcaa tgaaagctga actaatatgt 3360
tcatgaccaa taggcgtatg agtaatatta gtagactaaa tgtttaaaaa cgctaaaggt 3420
gtaatatcaa aattaatcta acaaaatcat tcatggcttt ggttgttgac tttctggagt 3480
gtaataataa tacaaattta aattttcgta aaaacaattt gatccgtacg tacacgcaca 3540
aatctaacaa actattgaaa tggaatcaaa tagctgaaat aatatttata aatgaatata 3600
tatttttagg acttaagttc actgaaggct tagcgggaaa taagttttta aggatgaaat 3660
attaatgcta ttatgctaaa ctctgttgag tttccagatg catatatcct ttgattttga 3720
actttcaacg aattttccac ttgatcaatc gaccttttgg aaacgaaaca agacccatgc 3780
aattaaacga catgcgttga tgccatttta agcgcttgca cttgcagtga attatgatta 3840
cgcgtggaaa acataattaa tgcataatta atttcatttt aaaatcagca ttaaacgcca 3900
gttaattaat aacagtaaac gcaattactg catagcatag ccatcaaaag ccataaaaaa 3960
gacagtaaat tgaattacaa cgtcccgtca atggtcgtga atatttaagt aaacgcgaaa 4020
agctcgcaaa tgtttgcaac ttaataaatt atacttggga cgaataatga taaacaagag 4080
aacagccgca aatataaaaa gaaaagccaa gcgaaaaaag gcgcgaggca cgagaagctg 4140
gaaaattaat tgagcgcaaa ttggtgaagc tg 4172

<210> 95
<211> 1014
<212> DNA
<213> Drosophila

<400> 95
atggaagagg atgtgtatgc ctcgctaggt gcctacaacg acagcggtgg cgacgattgg 60
agcagctcgg agcacctggt cctgtgggag gaggatgaga cgcagcgaac gactgctaat 120
gccaccagtc ggcataatca actgcatgtg gccaggtgga atgccaccgg caatgcgacc 180
atcagcgcga ccttcgagga cgtacccttc gatgcgaaca actactgggc cctattggcc 240
ctcgtcctcg tcctgggaac agcggccggg aacatccttg tctgcctggc catcgcctgg 300
gagcgccggc tgcagaatgt gaccaactac ttccttatgt cgctggccat caccgatctc 360
atggtcgccg tgctggtcat gccgctgggc atccttacgc tggtgaaagg tgagtccttt 420
gtcccattca aagtccggaa cagagtggaa gagtataata atcaaagcca gcgacagaaa 480
cgtggcaaag acattcaggc caggacctcc tgcgatgtcg ccaacacgag gcacgagttg 540
catttaatga agctgctggt gggaatggcg ttggcggtgt gcagattgca ttatgtctca 600
gatatttcga tcaaagctgg caaacatgtt gcaacactga tgacagtttt gtttgtgcgc 660
tttgggagcc attacaaaca caaacacaaa cacaaacacg agcacgggca ccaatacaag 720
cagaaagggc gacgcacaca cacgcaaaca cacacctgga gcaatccttg gcaaagtggg 780
caatgtagca agaaggagga ccaacaaatg gctgcttgtt atatcaagtg tgaaaaatat 840
tgcagtggca gtgacaggag cagttgttgc tgtagttgca cgggcacacg cccgccaact 900
ggaggcaaga aaggacgaaa tcgggatttg cctgctgtca agggctttgc agagggtcct 960
gcgacagaag ctgaaatctc cagcgaacag caggcattag atcccaagga ttaa 1014

<210> 96
<211> 337
<212> PRT
<213> Drosophila

<400> 96
Met Glu Glu Asp Val Tyr Ala Ser Leu Gly Ala Tyr Asn Asp Ser Gly
1 5 10 15
Gly Asp Asp Trp Ser Ser Ser Glu His Leu Val Leu Trp Glu Glu Asp
20 25 30
Glu Thr Gln Arg Thr Thr Ala Asn Ala Thr Ser Arg His Asn Gln Leu
35 40 45
His Val Ala Arg Trp Asn Ala Thr Gly Asn Ala Thr Ile Ser Ala Thr
50 55 60
Phe Glu Asp Val Pro Phe Asp Ala Asn Asn Tyr Trp Ala Leu Leu Ala
65 70 75 80
Leu Val Leu Val Leu Gly Thr Ala Ala Gly Asn Ile Leu Val Cys Leu
85 90 95
Ala Ile Ala Trp Glu Arg Arg Leu Gln Asn Val Thr Asn Tyr Phe Leu
100 105 110
Met Ser Leu Ala Ile Thr Asp Leu Met Val Ala Val Leu Val Met Pro
115 120 125
Leu Gly Ile Leu Thr Leu Val Lys Gly Glu Ser Phe Val Pro Phe Lys
130 135 140
Val Arg Asn Arg Val Glu Glu Tyr Asn Asn Gln Ser Gln Arg Gln Lys
145 150 155 160
Arg Gly Lys Asp Ile Gln Ala Arg Thr Ser Cys Asp Val Ala Asn Thr
165 170 175
Arg His Glu Leu His Leu Met Lys Leu Leu Val Gly Met Ala Leu Ala
180 185 190
Val Cys Arg Leu His Tyr Val Ser Asp Ile Ser Ile Lys Ala Gly Lys
195 200 205
His Val Ala Thr Leu Met Thr Val Leu Phe Val Arg Phe Gly Ser His
210 215 220
Tyr Lys His Lys His Lys His Lys His Glu His Gly His Gln Tyr Lys
225 230 235 240
Gln Lys Gly Arg Arg Thr His Thr Gln Thr His Thr Trp Ser Asn Pro
245 250 255
Trp Gln Ser Gly Gln Cys Ser Lys Lys Glu Asp Gln Gln Met Ala Ala
260 265 270
Cys Tyr Ile Lys Cys Glu Lys Tyr Cys Ser Gly Ser Asp Arg Ser Ser
275 280 285
Cys Cys Cys Ser Cys Thr Gly Thr Arg Pro Pro Thr Gly Gly Lys Lys
290 295 300
Gly Arg Asn Arg Asp Leu Pro Ala Val Lys Gly Phe Ala Glu Gly Pro
305 310 315 320
Ala Thr Glu Ala Glu Ile Ser Ser Glu Gln Gln Ala Leu Asp Pro Lys
325 330 335
Asp

<210> 97
<211> 9268
<212> DNA
<213> Drosophila

<400> 97
tataaatgaa ttctgaaaaa tgtataaaat gaaatggact catcattagg tcaatttgtt 60
tggaatcaaa aatttcccaa agaaatggtt ttacatttca gaaagtgacc cttaattttc 120
gccatcaata tttattagaa cattgacctt aaagcaattt gtatagtcgt cactttaaca 180
gatcaaagtt atacttaaag tcgtgcagtt gcaactgaaa ctgattcatt ctacgcctta 240
tcgatccatc ttccccgaa cttatcccaa tttctgttgg ctatctcgtt aggaaattcc 300
ttttcccagc ggccatacgg attcatgttc gaccagattc tggattcaca aatcccgcgc 360
ccaaatctca gggcacacgc aacaacacag taaacacaag tatcactcac cgaatgtgtg 420

```
ggagattctc ggccacacaa gctacagaga aacaaaccca attcaaacta aaaatttcta 480
aatggaaaat aaattaattt tattagttct ttataagata ctaagactta tactagttgg 540
aaatcttcta ggcatggttt gtaaatcatc cttatatctt tcgatattat tctaagactt 600
ttttcgtagt gcacacacaa ttgtattttt atgggaatag ggcgagagtg gtgcgctgag 660
ttatgacgtc caacagctgt gcaaaacgat ggccacagca tcgggaatag ccgctcatat 720
gtgggccagt tggcctgaaa tatgttgcat aagctccccc ctcttggcaa tatcatttgg 780
cagctcacga atgaactgtg gctcccggga aaatatggcc aaaaaaaata aaaaaaaaaa 840
ggcggggagt attaaataaa aatggctaga gaacattgca ttgtttgccg ccgaccaagc 900
caatcgtaaa tcgtgtgtac acataggtag aaagtattgc agccacttaa tacagtcata 960
agtcggccat ggaacattta caagctgacc aagaaataca atgaagaatc ggtttttttt 1020
ttcttttaca ggctactttc ctttgggctc ggagcactgc ctcacctgga tctgcctgga 1080
tgtactcttc tgcacggcca gcatcatgca cctgtgcacc atctccgtgg acagatattt 1140
atcgctacga tacccaatgc gatttggccg gaataaaaca cgtcggcggg tcactctcaa 1200
aattgttttc gtttggcttc tgagcatcgc catgagcttg cccttgagct tgatgtactc 1260
caaggtgagt aataatgcta ttcttttgtt tatattctat attccaacaa tgcgtcaaaa 1320
tatagtaccc ttttggcatt ctatttccct taatctaata gtgaaaaccc attaaattcc 1380
ctagttaacc tagtactcaa aatctaaagt gaaaattaca aacactgtgc caagtcaata 1440
attaaaatcc aaagcgagac aagtgtctga aatattttgc gcatacatat gtcctttact 1500
tacccaaatt aattccgtgc tcgccctgct atctttcctt tgattccaga atcacgcctc 1560
ggtgctggtg aatggaactt gccagatacc ggatccggtg tacaagctgg tcggctccat 1620
tgtatgcttc tacattccac tgggcgtgat gctgctgaca tattgcctga ccgtccgact 1680
tttggcccgg cagcgccaga atctgggtgg tggacagcag acggcggcgg ccactcccgg 1740
atgggccagt ggatggctcg gccaggcacc ggccttgggt aaggatgcta gacaccacac 1800
ttctgacccc cactctcacc cagaagtgtc cacgaagagc tgctgcttct gctgcagctt 1860
aatatgattt tcagtttgcc taaaatttgc attttgttct ccggattgca ctttttgcac 1920
tgaaaacttt acacgacact cgacaaaagt agagagtagg aaatgggcgg atcaggatca 1980
gcaaggggaa agtgctcaaa agtggggagg ggtgacaaca ctagacaaac aatgcagaga 2040
catgtgccgg atactccccg aatccgtgtc cctgggcttt ctggagccca aagtgcgact 2100
aagccgcagc aatgcacaca aaatcgacat tgtcgtcgac attgaaccgg atggaatggg 2160
gcgggggggg gggggggcca acaccccagg acaatcaccc attccaccca cttgcacaca 2220
caaccttaca aataggatga aaatccagcc catatgtgta gatagctaat ttccacgttg 2280
ccatgaatcg agggccagct acgtaaatgg cttgtaatgc tctaagagct cttccatagt 2340
gctcgactag ttttaaaacc gaagttttaa gcaccaactt ttcaacgaag cctcgttaca 2400
gccccgttgc gtagggtagt ttttaaaaac tcaagagctg cgttaacttt tagttttatt 2460
aaccactttt tgaaatttcc ataacacttt tcagcgcatc tacaaaaaag aggaggaatc 2520
tttaaacatg tagctaataa ataaaaagtg aacagtctgt tattttagat ttaaagagtc 2580
taaattgctt cctgcatccc taattacccc tctcgttact agcccagtta actggagctt 2640
ttggctaaaa acaatccaag atccttcgct tccattatcc cttcacaatt caatagacca 2700
gattcaattt attttaagca taactaagtg gctcctaatg ccattgtctg cccgcagaat 2760
ttattcagcg ctcccatttc tcatccaatt ggattgcctt gagccccaaa ggggacagag 2820
agtaattttc tgggggacaa gctgtggatc tctggatgtt gaaagtcaat acagttgggg 2880
aatatttaat ggaatgcttc tggtggtggc gaatggtgaa tgctgaatgg taaatgctga 2940
tggtgatggt gactgcagtc gaatcgcgtg cactgcaggg catttttgcg aaaatcgtag 3000
cgtgcattta ggggcgattc gcaggacgcg gtgtccccga ttcgtctaag tggatttgaa 3060
gcctcagctg caggcattag ggtgtgtatg tatgtatgtg tgggtgcgca tctgtgagcg 3120
gccaagtgcg actgtgttag tgccgtggat ttgggcagtt attttaagcc aaatctctgc 3180
tgtaaagcgt ccgcatttcg gtcaggcacg tagtcggatt tgccgtattg tactttttac 3240
acgcagagaa aaatcttcgc atttattaca caaagaatca atgtgtctga ttaatataag 3300
caagatattt aatgcaagtt agaatagctt caagcactat ttcataaggt ctgaactgga 3360
ctaggagaag aatttctctc agtgcctatg ttgaaggggc atttaagtgt gctcgtgcga 3420
gtttattagc gaatttgtga cttttcgtag cgcactattg gactgaatgc aatttaatgg 3480
aaactgtata tccgtttaac cccgtgttcg tgctccacag aacgacgatg cacctggcgg 3540
cgccttttga aaccgggtcc aggtaacgcc tcctcggtgc tgcacgcaca ctcggccaat 3600
tcaacggata cggatctcag cactttggac aaccacgagc tctggctgcc cgattccagg 3660
tgagctatcg agtttgcgag ccagtgttgc agtgcaaatt accatatata ttattcatgt 3720
ttatgtagca taaaggagcc aacgcccacg accatgacgg ccttgcatca gttcggtgcc 3780
gagatgctga aactatcccg aggcctggag tcggtggcct cgtcatccac caccggctca 3840
cccaccaaat ccgaattttc cctatcgaac cacctgcagc tgcagtaccc cagcagccca 3900
cagcgatatg ccgccaccca ccagcagtcc caccaccaca accaccagca gcatccggcc 3960
ggagtttacc accagcaggc atcgcccaaa ggtcgccacg gaaccacagt tctgggtctg 4020
```

tccaccacca cgctgggcat cgagcgggag agcaccagga attccctggc cagcagtcgc 4080
atgggcggcg aacagagcga tgggtgagtt cataaatagt aatataaata agtgagtcca 4140
aataatgggg tctaaaaata tactatctaa aagtctggaa gccaaccaaa attatggtta 4200
caattttctc aagatgtaat atgatgctca aacataatta aattgcttca ccacttaacc 4260
agaatctctg ttgtcacaaa taatattaac acagacaatt aaattctttt aactttatcc 4320
cttttgcggg ccacaatttt cactagaatt gtgtcaaatc ccttgaaaag catagacaat 4380
tcgggcaggt accttggaaa aacgtaaagt tttcgtggct cttcacctga caaaaggaca 4440
aacagctgcg atgggcaaac cgttttgcca gagtgatcct ttctagtcgc tcgcaggacc 4500
ttttgttgtc gctgtctcag ccggttttcg agtgccagta ctgcgcttgg ccaagtgtgc 4560
tgaccacacc aaaagccaag aaaatgctca cctcgagtcc tttttgtgac tcgtttactt 4620
tctccttttc catttattta gtttctttt ttcgttagct catttttgtg gcattgcctt 4680
ttattgtcct tccgccttcg acttgttgtg aaatttttcg taataattaa tttttttcg 4740
ctacgttatc ctttttcttt gtgtcggtgc aaactgtttt atttcttttt tttggccaaa 4800
aggcagaggc actgcctttt tatttggcat tcgagccgct caagtgattt tggcgtcgac 4860
tgaggccttt tgccaaaatg tgtttttcat taatttccta aatttagtct actggcttat 4920
gctgggataa gtgatgcgac gagaaattta attaaaaaat tgagataata gaagacgaat 4980
atattttgaa attaccgaga atttaataag aactggaaac atggctaaac tggcatattc 5040
aagtaaatag ggcgcttcat gtttatgggc attcaggtgg aactaatatt aactatttaa 5100
ttgcgtatga tatgaaattc tatggaagag cctggcaaat tctaaaaata cttttatttt 5160
aggcactcaa cccaattcga tattgtacac gcattttatt ttgatttaat tggttatcat 5220
atttagtctg caaaatccat ccaataaaag catgaatttt tcactgcaac tgcgaacatt 5280
gatatcttga atcccccgtt cgattttaag cttgaataaa aggcagctca acttgcgaaa 5340
tgactgcaca ttttatctcg ctcatatttt tcccatttaa ccaatccctt atgcatgttt 5400
atatcatata tcattgtgac tgtgcgcctt tcagccaaat gcttggtatt gggtattggg 5460
agcactatgt gccacattgt ttgacagttg aacgaatttg cacaatcaat aaaatggatt 5520
tggctttggc tccggcttcc tattgtccgc aatggatggg atgccacttg gcataaagga 5580
cgaaaagaaa cgggacgagg tcaaagtcga gtggttttta attgttgcac aatttgccgt 5640
ctcatgcggc aaagttgaca aatgatgctg gcgatattct aatttttgaa cttggctggg 5700
agacttaaaa tcagctcaac acttcgccac ataagagtat atattatatg tatgtgtatg 5760
cgaaggcggt atcctgcgat gtcctgatga gcttggcagc tgtttaattc ctttgtgagg 5820
gggaaacaat aacaaagttg gcatcagctc ggctttgttt cacttgcata atgagtagct 5880
ggccccaaat accacctccc cactccccca ttgcttagcc ctgttccaac gcatctcaat 5940
ttatgcgcac gtctcgtgca aacaaatcaa aaacgcgaag gatgcataat gagctgacat 6000
ccttatcctg cgtcttggca gcaccgagca tcccgatccc tgtcctccct gcagccacca 6060
ccacttggcc ttctgtctgc taacgccttc attagcagct aaccacgccc ccaacgccac 6120
cgccccacac cgtcttggcc aaataaagtt ggcgcgcatt tgcatgcagg gcaaaacggg 6180
cggcaggaca gaaaaccgca aaggcagctc gcaaatttat gcgctttatt aaaatgaatg 6240
caaaatgagg cggtacagaa aaaaaatcga tgtacaatat gtttatatat aactttaaat 6300
taaagctaca catatgagct aagtatttct gcaattaata atataaaatt aaaaaaaaaa 6360
aaaatcccgc cagtaaagta cagtaagccg aactaacata ccaaatataa tatagtaaat 6420
aataatgcaa gaatttttca agtgtaccac aaacaacgga gcaagaaaaa atgttagaag 6480
gcggtaaacc acaaagagca aacacaaaaa atgttgccta cttcccggca aaacaggtga 6540
aagaaacgag ctgaccacaa agcgtgctgg cttttgttca attgccgtta gcgagtgttc 6600
actgcattgt tttaggcacc tcaccttcga atagagtttt gccaattggc tgcgcaattt 6660
taattaaaaa tgtacgacgc cgaaggggag aaggtgaaag aaaatggctc ggcattgggt 6720
actaaatcaa taaatcggtg gggtatgaca tagacatatt ttatatattt tcattattaa 6780
aaccacatgg ttccttttgt tctagtctgc tttaatttga ttcaggaaaa tcgtttaata 6840
gaccactttt ttaaaatgaa atgaaaataa tcatttaagg actttgactg tagtatgtca 6900
ttatatttat ttaatcagta aagttctttt cctttcagga ctctttcaca gctttcccaa 6960
cgattgcgag cgtataaaaa gcgacgaaga gcttcgtcag ctgttcctgg acgtgagaaa 7020
cgtgccgggc atgaggacgc agacgaggat gtggaaacgc ccacgtcaac gctgcgccgc 7080
cacaagcgcc ataacagtct gcccaagaac gcactctatc cgcgacacac cactgtacag 7140
gagagcctcg atgacgacga cgacgaggag gaggagagcg cgcaacagaa ggatcagcag 7200
acgaagtcga agatctgcca ctcggacact gagctagatc ctccgcagaa ggcgaagcag 7260
tgccacatag ccgggcagaa tcatcagtcg gactacctgc agctgccctc ggtttgcact 7320
tgtccgtatt ttggggacag accgctccag aattgtgtga aatcggcgga ggtgaagatt 7380
atttcgtccg ccttccgagt gaccaccacc acgacggcgg tgagcagttc gcccagcgag 7440
atggagctac tcatgtgcag tagtggcaat aaaaagtcat tgacctccag tgtgagtgcc 7500
ggcataacag gtggaccggc aggagcggct ccagccacag ttggctccac cctcagtcca 7560
cactctgctc acaatcaggg cagctcgctg acagttcaga gcgatggtag tgggtacttg 7620

```
gctgctccgg gaactccttg tcctggccga cgaaagttga gcatctcaaa gaccgcctcg 7680
gtggtcacgt gggattcaag tcgtcatcgt cggcggggca gcagttttgg cggagtgcga 7740
acatccctgc tgttgacgcc caccaaaaca gcgaccacct ccaactcatc cacaccgttg 7800
agaagatcgg ccaccttgag gagtcaccag aatatgaact accagggagc gggcgaatgt 7860
ggtggcaaga ccaggacgac aacatcctca ccctgcatgt tgcagcgtca gcagacggtg 7920
cgatctcatc actcgcggaa ctccagcgtg atctccagaa actcctcgcg ccacggcagg 7980
atcatcaggc tggagcagaa ggccaccaag gtgctgggag tggtgttctt tacgttcgta 8040
atcctctggt cgccattctt cgtcctcaat ctcctgccca cggtgtgcgc ggagtgcgag 8100
gaacgaatta gccactgggt cttcgacgtg gtcacctggc tgggctacgc tagctccatg 8160
gtgaacccca tcttctacac catcttcaac aaggtgttcc ggcaggcgtt caagaaggtc 8220
ctgctgtgcc gctattcgag cacgagtgcc tggcgaccga gcagataaca gacgcccgtt 8280
aaaccgggca aactgaagac gaatacccac aagtgcgcca gtgtggattt tcaggattcc 8340
actgctccgg tgcggggcgt tctgaaaaga agtgtggcag tggcaagtga catataaata 8400
cggcaaaggc cggtcatttc gcctgcggtc caccgcacaa cgtgtcgtat acgcaatctc 8460
gaaaggacga taatgcccac ccgcactcca caccacgcac tccgcactgc tcacacgtgt 8520
gtccttgggg cttaaattct tatgtgcctc taattaattt caaattcaat ttgcatgcca 8580
ttgaaaacac atgtcccagg cgtaagaagt ggggccaagt gaattagata ttgaagtgga 8640
gtcggatccc cgcaggattg ccaagggtcg gggattaaac tagccacaga caacagcata 8700
acgagcatgc tactcacccc tggtggttcc ttccatttcc ctcagctctc agcaaacggt 8760
gaagatgcac ttaattgcgg aatgtttgct ttcgagcaac gtaataccgg caaacacatt 8820
tactttatag caacgctctt cagttatcga caaagttctg gctcaccgaa agttcatgtc 8880
aatctgttga tatttgatag cattttccgc gaaattcctt ttagattcac tacgtactcg 8940
aaaataaact tgaaaattca cgtaaattta aagattgtat atataaagtt caacttttga 9000
aattcttaaa aataaagcgc caatgtcgct tcgcaaactt cgatcgataa acgctatttg 9060
gtttggcatt tggcaattat gcattaaatt tatggcaaat tcactgcccg ccaatgcaca 9120
aagaatttca tcggtatcca ttaaagaaac cgacacgtac acaaatccac ggcaatttgt 9180
gacttttgtt gtttgtttgt gtggttggac cggcagattg ctatttttat ttgaaaagtt 9240
cattcatttt gttttatttc tgtggcca                                   9268
```

<210> 98
<211> 1251
<212> DNA
<213> Drosophila

<400> 98

```
atgaagaatc ggtttttttt ttcttttaca ggctactttc ctttgggctc ggagcactgc 60
ctcacctgga tctgcctgga tgtactcttc tgcacggcca gcatcatgca cctgtgcacc 120
atctccgtgg acagatattt atcgctacga tacccaatgc gatttggccg gaataaaaca 180
cgtcggcggg tcactctcaa aattgttttc gtttggcttc tgagcatcgc catgagcttg 240
cccttgagct tgatgtactc caagaatcac gcctcggtgc tggtgaatgg aacttgccag 300
ataccggatc cggtgtacaa gctggtcggc tccattgtat gcttctacat tccactgggc 360
gtgatgctgc tgacatattg cctgaccgtc cgacttttgg cccggcagcg ccagaatctg 420
ggtggtggac agcagacggc ggcggccact cccggatggg ccagtggatg gctcggccag 480
gcaccggcct tgggtggacc ggcaggagcg gctccagcca cagttggctc caccctcagt 540
ccacactctg ctcacaatca gggcagctcg ctgacagttc agagcgatgg tagtgggtac 600
ttggctgctc cgggaactcc ttgtcctggc cgacgaaagt tgagcatctc aaagaccgcc 660
tcggtggtca cgtgggattc aagtcgtcat cgtcggcggg gcagcagttt tggcggagtg 720
cgaacatccc tgctgttgac gcccaccaaa acagcgacca cctccaactc atccacaccg 780
ttgagaagat cggccacctt gaggagtcac cagaatatga actaccaggg agcgggcgaa 840
tgtggtggca agaccaggac gacaacatcc tcaccctgca tgttgcagcg tcagcagacg 900
gtgcgatctc atcactcgcg gaactccagc gtgatctcca gaaactcctc gcgccacggc 960
aggatcatca ggctggagca gaaggccacc aaggtgctgg gagtggtgtt ctttacgttc 1020
gtaatcctct ggtcgccatt cttcgtcctc aatctcctgc ccacggtgtg cgcggagtgc 1080
gaggaacgaa ttagccactg ggtcttcgac gtggtcacct ggctgggcta cgctagctcc 1140
atggtgaacc ccatcttcta caccatcttc aacaaggtgt tccggcaggc gttcaagaag 1200
gtcctgctgt gccgctattc gagcacgagt gcctggcgac cgagcagata a          1251
```

<210> 99

```
<211> 416
<212> PRT
<213> Drosophila

<400> 99
Met Lys Asn Arg Phe Phe Phe Ser Phe Thr Gly Tyr Phe Pro Leu Gly
 1               5                  10                  15
Ser Glu His Cys Leu Thr Trp Ile Cys Leu Asp Val Leu Phe Cys Thr
            20                  25                  30
Ala Ser Ile Met His Leu Cys Thr Ile Ser Val Asp Arg Tyr Leu Ser
        35                  40                  45
Leu Arg Tyr Pro Met Arg Phe Gly Arg Asn Lys Thr Arg Arg Arg Val
    50                  55                  60
Thr Leu Lys Ile Val Phe Val Trp Leu Leu Ser Ile Ala Met Ser Leu
65                  70                  75                  80
Pro Leu Ser Leu Met Tyr Ser Lys Asn His Ala Ser Val Leu Val Asn
                85                  90                  95
Gly Thr Cys Gln Ile Pro Asp Pro Val Tyr Lys Leu Val Gly Ser Ile
            100                 105                 110
Val Cys Phe Tyr Ile Pro Leu Gly Val Met Leu Leu Thr Tyr Cys Leu
        115                 120                 125
Thr Val Arg Leu Leu Ala Arg Gln Arg Gln Asn Leu Gly Gly Gly Gln
    130                 135                 140
Gln Thr Ala Ala Ala Thr Pro Gly Trp Ala Ser Gly Trp Leu Gly Gln
145                 150                 155                 160
Ala Pro Ala Leu Gly Gly Pro Ala Gly Ala Ala Pro Ala Thr Val Gly
                165                 170                 175
Ser Thr Leu Ser Pro His Ser Ala His Asn Gln Gly Ser Ser Leu Thr
            180                 185                 190
Val Gln Ser Asp Gly Ser Gly Tyr Leu Ala Ala Pro Gly Thr Pro Cys
        195                 200                 205
Pro Gly Arg Arg Lys Leu Ser Ile Ser Lys Thr Ala Ser Val Val Thr
    210                 215                 220
Trp Asp Ser Ser Arg His Arg Arg Arg Gly Ser Ser Phe Gly Gly Val
225                 230                 235                 240
Arg Thr Ser Leu Leu Leu Thr Pro Thr Lys Thr Ala Thr Thr Ser Asn
                245                 250                 255
Ser Ser Thr Pro Leu Arg Arg Ser Ala Thr Leu Arg Ser His Gln Asn
            260                 265                 270
Met Asn Tyr Gln Gly Ala Gly Glu Cys Gly Gly Lys Thr Arg Thr Thr
        275                 280                 285
Thr Ser Ser Pro Cys Met Leu Gln Arg Gln Gln Thr Val Arg Ser His
    290                 295                 300
His Ser Arg Asn Ser Ser Val Ile Ser Arg Asn Ser Ser Arg His Gly
305                 310                 315                 320
Arg Ile Ile Arg Leu Glu Gln Lys Ala Thr Lys Val Leu Gly Val Val
                325                 330                 335
Phe Phe Thr Phe Val Ile Leu Trp Ser Pro Phe Phe Val Leu Asn Leu
            340                 345                 350
Leu Pro Thr Val Cys Ala Glu Cys Glu Glu Arg Ile Ser His Trp Val
        355                 360                 365
Phe Asp Val Val Thr Trp Leu Gly Tyr Ala Ser Ser Met Val Asn Pro
    370                 375                 380
Ile Phe Tyr Thr Ile Phe Asn Lys Val Phe Arg Gln Ala Phe Lys Lys
385                 390                 395                 400
Val Leu Leu Cys Arg Tyr Ser Ser Thr Ser Ala Trp Arg Pro Ser Arg
                405                 410                 415

<210> 100
```

<211> 8877
<212> DNA
<213> Drosophila

<400> 100
gacctttaag cagactaagg ggttgaggca atcctttttt cccaactcgt gtcaattaac 60
tcattaaact tcccaccgtt ttgtggagcg tcggacccaa ttagctttca attataccgt 120
ctatgggctg gtttttcata atcaagcaaa atttcaatta attaaaatgt aaccgagcta 180
aaatggaggt gtgagaaaga gcctttatta cgaaggactc tattacacat taaagtttta 240
tggaaagaat gaatattctc agtgagcaag accagataag ctgccaacat attttctcta 300
tcctgtcagc ttcataaaat gttggttttg atgatggaaa cttttgccca tgaaactttt 360
tgttcgcaga ttatcgtgaa agagttttga gctccgtttc gctgtattgt tggcttgcca 420
aactgttgca tacttttagg aagttgcaag gagcggaaca tgtgtttgcc atgtcaactt 480
tttgccatat ttgtagccat aacagctaca catctccgag gaaggtattt taggaatttt 540
ttttgagtac atcaaagccg ataaagaaaa tcccccagac aatggaagct taccgaaaga 600
aaaccgcaat gataattccc cagagaaaaa cgccgatcga gtacactgaa gtaaatagac 660
catttattta gatgctagaa tttcttgttt atacagagaa ttaatccatt tgcacattga 720
gttacattca acagctacag ttggagttca tagcatggag ttggttggat tgttaatggg 780
gctctcagct gctcacatat acgacattta acagtttaaa acgattgacc taacttatgt 840
tgcttgtttg ctccgcttgg tacttggttt cgtcctttca tcgcactcca gcctttaata 900
attagtacta gctaactaat cgcactcgca gatctgcgac acggaagtgg aaagttcaat 960
gaaagttcag tctcttaaga tgcacttgct agggatttgt ctacaccgag ttctcctcga 1020
gtccttgctg gaattcgaga atgattattg gatggtctgt atgagtgaga agagctgaat 1080
tctggttctc accctcatca ggggactaat ggcccggggc attacgctca gagtgctgcc 1140
ggaactcatg gcatttgccg ccaaagtggg catcgtgtgc agggcattgt tggagccccg 1200
gtggagtgaa agtggggcat gctgacccac cagcgtggtg gtggtggccg aacttacgca 1260
cgattcccgc ttgccacgcc catagtagct ggtcaaaggt ctgcaacgga gtggatgtca 1320
agatatcccg tcacttctta gccacgattg ccacataact taccgcatac tctcggctcg 1380
gggggaacca ccgcctttcg agaagctttt cgtggtataa ctgccaagtt tacgaagaaa 1440
tgcgcagaaa aagttagtgt acaatagtga aaatattttt aatagctgcc attcctagta 1500
gcaatctagc attgccacac tttatgatag gtttaacatt aaaggcattt gaggagaaat 1560
atgtacaact tagttgtttc ctatataagt ttgcagtttt ctcttctcga gtagtgtatt 1620
tattcttaag tgtacttaca agcagactta aaagtttgga ctacttaccg tcggttctgg 1680
ttaagctgaa tggtgcgcgt gtcccgccac gtggaaatgt ggtgtctcag tgcatttcgc 1740
acctccgagt tgaggaaaca gtagaacagc gacaccgaga atccctgtgg cacacacaca 1800
aacagatctt tagaaaaaaa gtggctaaga ggattgacaa acgaaaccca cctgagtgct 1860
gagtaatacc gctcgcagga cggcaaacat gtgtcccatt aggccagatt ccgagggacc 1920
agccagcacc accaggtagg tgatgccaaa aagtgggatt agcaccagca gtgccttagc 1980
agctttccta tactgacgag tctccactgt attggccgag cgcagctttg tgattagaac 2040
ctaaaagaga agcaacccga taatctggct cagtattttc ggcaatttac agtcggaaaa 2100
atgtcggatg ataaattgta aattcataca atagctttaa agtgtttcaa gaaactaatt 2160
gggtgattca taaatgtgag tttttgcatc catgttaact tgttcaatct tctctgatgt 2220
tcttttttc gcatatcaat attttttctg tgtaaattcc atacgcaccc acatgatacg 2280
cagcagaaat gtgagattga ttattagcac cgcacagaca ggtccctggt atatccaatc 2340
cacatgggtc tcctgcatcc aggggcagtt gatttcgtac tgtcggatgg gaaacaaaag 2400
gaagactact ctccaggggt ctcttcaaat cccaaaatag acataccttt tcgggggtgc 2460
tgtaggtgac cgtcagactc ttggccacag cccaggtgac aacaaacaag gccgggccac 2520
ctgccagcag gagaaagact taattggctg gcacacacaa tcgggcattc acacttaccc 2580
cagcctatgg aggcgtagat attaaagcgt aaattgtccc cggagaaagt cttaaccacc 2640
agcatgtaca ggtacaaacc ctcgaccagc atccaaaaga agttggtcag ggtgaagaag 2700
tggaacaggg tgatcagagc aatgcagctg cccactccac ttcggatgga gatctgtaga 2760
ttaagtggta gccattagtc ttaagctgtc agcgtctgtt tatagtttcc tttcactttg 2820
gactagcacc cctttcaggc acgagtatga tttaataagc cagcaattgc aattatctgg 2880
cagtctcata atgctcacct ggacggataa caagagtatc cagaacaaag ccgacatgat 2940
gtacgtgaag aacaagttgg cgtggatggt gttgcgcagg caacgcagct ccctggtgaa 3000
caaagacaaa tacaaaggac gggggtaag tcgggggtta ggcaaaaata aatgccatac 3060
ccagggtatc ggggtcacca gggagggggc tggggtgggt acactcactt gaagtaggcg 3120
aaaacaatca gcgccagcga gagcgatacc aggctgaggg tgtatccaat gtagtagata 3180
atggtgggca gctccacgat gacctcgaac tccggcacgg attctggggc gggcaggtgg 3240
gcgcaggcat cgtagttggt gtacttctcc caagttccat tggcgtggca gaacctcgtt 3300

```
gcattctctg gtcaaataca aaaacgcacg cagcgaaaaa tgcaaacagt ttttaaatag 3360
actagtccta tacacttgga tctatattgg gaatatgcct tctctgtttg gacatcgact 3420
taaaaaattc caatttaaat tcataactta gtgcaaagct aacagttttt ttcaaacgct 3480
tagtttaaat ataagtctaa tgttaaagta taaggtcaaa ggggtactca ctgctgctat 3540
cgtaatgaat tccctgcaac tcgtccatgc actggagcac cgccagggta ccgcgagctg 3600
tcctcggcca gcacaggatc gagtcgaact gggtgaggca gtggccacta tcgttgccat 3660
aggtggatgc ctcgatgtgc tcctgcacca ggcactgcaa ctccacgctc tcaccgatgc 3720
catctaggtt gtgcagatcc aggagcggat cactcccgct ggcgttcacc gaatcgatgt 3780
ggttgtggtc actcattcta caacatcctg gcacttggca tctctggtac tgaaataagt 3840
cagtttaaag ttgaggcaaa gattgtaaaa tacttgaagc agtgggaaaa tttaaaatat 3900
tttagaatca ttatagagtt ttaatacatt tcatttaaga accgtcggac aataacgttt 3960
tagaaaaatg gtaagcttta atcggaaata aaatattata gactgatacg aaaatattgc 4020
tgtcatttcc tgaccatttc cctacgaatt gacaacgttc ataatactcc ccagctgaat 4080
tcgttcaatc agaaagttat aatttcctca attgttaggg gattccccgg cggttggctg 4140
aaactaagct gtcgggcaaa gttcaattgc caacccttct gctcgctcca ccgcagccag 4200
tggcatttcg cctgtgtgcg ttaattgcat tggccatgtt tcaccgacaa cgggccacgc 4260
ccacatcgaa ctcggcttaa agatccatca actcggcggg aagtatgctt taagccaaag 4320
accatttcta gggaaaacag ggtggctaga gataagtaaa tgcacaggcc cagtgccata 4380
aaaattcatt agttttgcca tcaaaggcgg acagaagcgt cagcaaatca gagaagcggc 4440
aggcaacctt ttattctgat atgtgcccat aaaaagcacg tcaaaggacc caagtctggg 4500
gtgatggtga tagaaattag aaaagctaac ttccggccaa agattgtggc ctttgtggcc 4560
tggttctgcc tggttctgcc tcagagacga aatctcgcat ttatgttctg ataataaatg 4620
gcataaatga tgcatcgtta cattaatgca cgacataaaa cggggcggca agtaagacac 4680
acacacaaac aatacatatt aaaaatgtag atgggcctgt ttgtacagtg agctagaaaa 4740
gtaaaagtgc gctctttaaa atttaagcgg aagtggcttt cactttctac ttaacagtga 4800
aataaattag gcaaacaagc cagcttgaat tccaaggaga agccataagc caaggcagga 4860
atttaaatcc actcggagtg aaaatggtta tctcttaagt gatgcccaac ccgaaccaat 4920
tactttctgt ggaatttccg attttctcgg acggcacatt ccccaggcga ttttctaatt 4980
ggcacaggaa tttctaacca caagccatga aatattaaac ccgaagggaa attgaaaatt 5040
ctgccccaag gcttgggtgc tccacttctg gctgccactg tatgctcggc ttcgtatttt 5100
tctttccatt gtttctgctg tttgtttttc tcgtaaatta acgcaatgca tgtacataaa 5160
acaaaggctg gctgattatt actgcacggc acatgcctgc acgtcctttc cggagtttgc 5220
caaggtcctg ggtcctggtt cattagccgc tcgcttactt ataacacata cgcccccgttg 5280
ccttcatttt ccagcttcgc acgctcacag caacgcagat gagttaagac agcaatatac 5340
gcgccgcaat tttaagttga aaagcatatt atttgcaaaa aagggaaaga acggcgaagg 5400
accgcctcta attacctcat aaataacgtc gcacgaaatt agctggccta aattatggcc 5460
ataactgcta attgcatatt tgatttcaat tagttcagct gacaggggcc caaaaaccca 5520
agaaggacac atttgaaatc agcgtgcgaa tgcaaatacg ttctgacagc tcgcttggac 5580
aaatctcatt tatggccagc caatttgacg aatgtccgcc cgccggggaa acaggcaaaa 5640
caaattcagc ttggccgaat catcgccggc ataaaaagga cattaccaga aatacataca 5700
tatacggagg cataagaaat ggttggtcct ggacgaagac caacggacaa atcgcacatc 5760
gtcctttctg ctcctctgcc tctttgcagc atgctattat agaagattgt atgtaatctc 5820
gtacgattgc acacaatgga tttgccaaag gcgacacaaa aaaaaaaaaa aaaaaaaaaa 5880
taggggattg cgtgggggga gtttgcttgt gggtccttta cttttttct cccccatgtg 5940
agaggttctg caatttcctg ctgctataat gcttataaat gttcacccgc aatcgccggg 6000
gatttcttta gggaattgaa aacataataa cagtgcgaat agtggcaagc attcgcattt 6060
tccacttatc caactaaata tgcagactaa aaattaaatt ccagtttttt caagccctcg 6120
aaggggattt ttattttacg cccttatatg tgtacctatg caaaaattcc gcaattccgc 6180
gggacaccgg ggaaaagtgc tattttcggt tctcattaat gtagctcatt aaattttata 6240
tgcataagct gccagctgca aacaaaaaat gtaagcattt aaaatttata agcgtgccac 6300
acaagagacc aacagcacat tacgtatacg cagcatggac aaagccagct tggaaagcca 6360
gcacgcaggc ataagcgtat agaaagtttg taattaaatt gctcatttaa gcgtggaaac 6420
aatgaattcg atcgggaata cagggagtgg ggagtgttct gaggaaatat gtaatattca 6480
gcgtgtaatc cgaagaaaat atattactaa ttttattcat ccgccttggg cgcttaaatt 6540
ttattgttcg gacttttgtg gactctcctc ggaggccgat aaggaaaagc cagtctagac 6600
aagccaacat cataatttgc aggcacgcac tgccatttcc atttccgccc aagcttgagc 6660
ctgatctgct ggccactttc catgcttaag tcagcacata tcctttagca ggaagatgga 6720
atgggatggg atggggctga cacttcagcg attgccaaca agtgttgcgt tttattgtct 6780
ggcctttaac agctctcctt tgccatccgc ccattgtgtg tgcggcacgt gcaggcgcct 6840
gcatttggag ctacgtgagg gatctccact cctgtggata gtccgagcat aaaaccataa 6900
```

aaatccgccg caaaaaaaga aacgaacctg gccaccaaat ctgttattta ttggcccaat 6960
aaaaatttgg tttaatgcat tacaagttgc cgaaatgcca cggagattgc cggagattta 7020
cttggcttgc aaagggaagt ccttaaggcc atggcaacaa gtccttcctg gttagcagtt 7080
agtttattag agtggtcaag tgcatggagt gggggcatt tcggatttcg ggctgaaatt 7140
ttcccaatcg atactgcaaa cacgcaaaac aatgcctctg ctgccaacta atatccatac 7200
tctcggctca ttcagccaca gagttttcgt tttagtaacc gctggcattg ttttcagcag 7260
gtgagcgtgt gtatgcactg gcattttaat taaccacaat ttgtttctgc attccaaacg 7320
ccctgaaaac gcttaccttc tactcctagc caagtttcta tctgcagaat gagattcctt 7380
tcgccactca cactcacaat ttgcgttgga attcttggga ttacttgcgt ttagcactcg 7440
gactacggac cgcaaaagcg ccaggaggcg ggtcctttgt cctttcgacc acctacttga 7500
cccgtcctaa atgcgtttaa cgccttccaa gctgcgaaat gattatgtta aagcggtttt 7560
tgctggagcg gttatttttg ctagttgcct tggttattga cacacgccca tttacgtggg 7620
caaaacttta ttgagttcac atttatatgc gggattattt gttatagcga cccttgaact 7680
tattcgaagc actcaaagcc aacgactaga aaagttattc gacattaatg cgcggcactt 7740
atttgttatt tcttttgcgt ttgctaatta aaattttcat ttttctcttc gcacattatt 7800
taaaaattag tttccacttt aacggctctg aggaggacaa taggttttcg caagttttga 7860
aattccgaaa cgaaccgcac tcgtacacat cccgctgata actgcgggga tacccgaagg 7920
atgtgtctta acaaggctcg ccttcttacc aaaaggcgcc aattctgggt tccacaaagt 7980
gaaaagctcg caaagctgca ttccacaagc tgttgaaaaa gctttccgcc gggcaacagg 8040
tggaacatcc tgacgggagg ttgggtaaaa ggaccccaaa cagaaaccga ttcgattgca 8100
cttctaatgc ctgaaaatct gcctatttga tttatgcaaa tatgtgccgg agtctcatgt 8160
ttgcccagca attcgctggc tgtttgttta gggcatgcgg caaaagggat accccggact 8220
caacaggatt cggcaggatt cgccaggata ttgtttgtga aactccttgc aggccgtctt 8280
tgaagtcgca acctgattta attgtgattt tcggagcccg ttcaaaacac ctctcacgtg 8340
ggcctaatca aattaaactg aagtggcagc gcctcgggtc tgacagttat tgaaatatca 8400
atatggctag gctccgggaa aatggtttcc cggaaaatgt gcgaaattaa attgtagaga 8460
aaacaagggc aacattttta ctcactggca tcttttctca tgcctgccaa attttatatc 8520
acaatgataa acatttagac ggatgtcctg gctgcttgga atgcgtgcga cacatgaggg 8580
ttgcctgtgt ggatttttgg gcgcaaggat atgggcgggg tggtggggc gatatctcca 8640
accaaactgt ttgattttta aacaaatgct aatgtcaact gccactcaga ggcatttaaa 8700
ttccaattta aaaatgcttt ctctcgatgc caaaagcaaa gacatgttcg ttgtcctgtt 8760
tcgagtccat ccactcgttc tttttctctc agattctccg attctctgac ccgtttctcc 8820
agttttctgc tcgcattaaa gcgttcgtcc tgtctcatgc atatttaatg tcacacg 8877

<210> 101
<211> 2089
<212> DNA
<213> Drosophila

<400> 101
cggttcgttt cggaatttca aaacttgcga aaacctattg tcctcctcag agccgttaaa 60
gtggaaacta atttttaaat aatgtgcgaa gagaaaaatg aaaattttaa ttagcaaacg 120
caaaagaaat aacaaataag tgccgcgcat taatgtcgaa taacttttct agtcgttggc 180
tttgagtgct tcgaataagt tcaagggtcg ctataacaaa taatcccgca tataaatgtg 240
aactcaataa agttttgccc acgtaaatgg gcgtgtgtca ataaccaagg caactagcaa 300
aaataaccgc tccagcaaaa accgctttaa cataatcatt tcgcagcttg gaaggcgtta 360
aacgcattta ggacgggtca agtaggtggt cgaaaggaca aaggacccgc ctcctggcgc 420
ttttgcggtc cgtagtccga gtgctaaacg caagtaatcc caagaattcc aacgcaaatt 480
gtgagtgtga gtggcgaaag gaatctcatt ctgcagatag aaacttggct aggagtagaa 540
gtaccagaga tgccaagtgc caggatgttg tagaatgagt gaccacaacc acatcgattc 600
ggtgaacgcc agcgggagtg atccgctcct ggatctgcac aacctagatg gcatcggtga 660
gagcgtggag ttgcagtgcc tggtgcagga gcacatcgag gcatccacct atggcaacga 720
tagtggccac tgcctcaccc agttcgactc gatcctgtgc tggccgagga cagctcgcgg 780
taccctggcg gtgctccagt gcatggacga gttgcaggga attcattacg atagcagcaa 840
gaatgcaacg aggttctgcc acgccaatgg aacttgggag aagtacacca actacgatgc 900
ctgcgcccac ctgcccgccc cagaatccgt gccggagttc gaggtcatcg tggagctgcc 960
caccattatc tactacattg gatacaccct cagcctggta tcgctctcgc tggcgctgat 1020
tgttttcgcc tacttcaagg agctgcgttg cctgcgcaac accatccacg ccaacttgtt 1080

cttcacgtac atcatgtcgg ctttgttctg gatactcttg ttatccgtcc agatctccat 1140
ccgaagtgga gtgggcagct gcattgctct gatcaccctg ttccacttct tcaccctgac 1200
caacttcttt tggatgctgg tcgagggttt gtacctgtac atgctggtgg ttaagacttt 1260
ctccggggac aatttacgct ttaatatcta cgcctccata ggctggggtg gcccggcctt 1320
gtttgttgtc acctgggctg tggccaagag tctgacggtc acctacagca cccccgaaaa 1380
gtacgaaatc aactgcccct ggatgcagga gacccatgtg gattggatat accagggacc 1440
tgtctgtgcg gtgctaataa tcaatctcac atttctgctg cgtatcatgt gggttctaat 1500
cacaaagctg cgctcggcca atacagtgga gactcgtcag tataggaaag ctgctaaggc 1560
actgctggtg ctaatcccac tttttggcat cacctacctg gtggtgctgg ctggtccctc 1620
ggaatctggc ctaatgggac acatgtttgc cgtcctgcga gcggtattac tcagcactca 1680
gggattctcg gtgtcgctgt tctactgttt cctcaactcg gaggtgcgaa atgcactgag 1740
acaccacatt tccacgtggc gggacacgcg caccattcag cttaaccaga accgacgtta 1800
taccacgaaa agcttctcga aaggcggtgg ttccccccga gccgagagta tgcgaccttt 1860
gaccagctac tatgggcgtg gcaagcggga atcgtgcgta agttcggcca ccaccaccac 1920
gctggtgggt cagcatgccc cactttcact ccaccggggc tccaacaatg ccctgcacac 1980
gatgcccact ttggcggcaa atgccatgag ttccggcagc actctgagcg taatgccccg 2040
ggccattagt cccctgatga ggcaaggact cgaggagaac tcggtgtag 2089

<210> 102
<211> 504
<212> PRT
<213> Drosophila

<400> 102
Met Ser Asp His Asn His Ile Asp Ser Val Asn Ala Ser Gly Ser Asp
1 5 10 15
Pro Leu Leu Asp Leu His Asn Leu Asp Gly Ile Gly Glu Ser Val Glu
20 25 30
Leu Gln Cys Leu Val Gln Glu His Ile Glu Ala Ser Thr Tyr Gly Asn
35 40 45
Asp Ser Gly His Cys Leu Thr Gln Phe Asp Ser Ile Leu Cys Trp Pro
50 55 60
Arg Thr Ala Arg Gly Thr Leu Ala Val Leu Gln Cys Met Asp Glu Leu
65 70 75 80
Gln Gly Ile His Tyr Asp Ser Ser Lys Asn Ala Thr Arg Phe Cys His
85 90 95
Ala Asn Gly Thr Trp Glu Lys Tyr Thr Asn Tyr Asp Ala Cys Ala His
100 105 110
Leu Pro Ala Pro Glu Ser Val Pro Glu Phe Glu Val Ile Val Glu Leu
115 120 125
Pro Thr Ile Ile Tyr Tyr Ile Gly Tyr Thr Leu Ser Leu Val Ser Leu
130 135 140
Ser Leu Ala Leu Ile Val Phe Ala Tyr Phe Lys Glu Leu Arg Cys Leu
145 150 155 160
Arg Asn Thr Ile His Ala Asn Leu Phe Phe Thr Tyr Ile Met Ser Ala
165 170 175
Leu Phe Trp Ile Leu Leu Leu Ser Val Gln Ile Ser Ile Arg Ser Gly
180 185 190
Val Gly Ser Cys Ile Ala Leu Ile Thr Leu Phe His Phe Phe Thr Leu
195 200 205
Thr Asn Phe Phe Trp Met Leu Val Glu Gly Leu Tyr Leu Tyr Met Leu
210 215 220
Val Val Lys Thr Phe Ser Gly Asp Asn Leu Arg Phe Asn Ile Tyr Ala
225 230 235 240
Ser Ile Gly Trp Gly Gly Pro Ala Leu Phe Val Val Thr Trp Ala Val
245 250 255
Ala Lys Ser Leu Thr Val Thr Tyr Ser Thr Pro Glu Lys Tyr Glu Ile
260 265 270
Asn Cys Pro Trp Met Gln Glu Thr His Val Asp Trp Ile Tyr Gln Gly
275 280 285

```
Pro Val Cys Ala Val Leu Ile Ile Asn Leu Thr Phe Leu Leu Arg Ile
    290                 295                 300
Met Trp Val Leu Ile Thr Lys Leu Arg Ser Ala Asn Thr Val Glu Thr
305                 310                 315                 320
Arg Gln Tyr Arg Lys Ala Ala Lys Ala Leu Leu Val Leu Ile Pro Leu
                325                 330                 335
Phe Gly Ile Thr Tyr Leu Val Val Leu Ala Gly Pro Ser Glu Ser Gly
            340                 345                 350
Leu Met Gly His Met Phe Ala Val Leu Arg Ala Val Leu Leu Ser Thr
        355                 360                 365
Gln Gly Phe Ser Val Ser Leu Phe Tyr Cys Phe Leu Asn Ser Glu Val
    370                 375                 380
Arg Asn Ala Leu Arg His His Ile Ser Thr Trp Arg Asp Thr Arg Thr
385                 390                 395                 400
Ile Gln Leu Asn Gln Asn Arg Arg Tyr Thr Thr Lys Ser Phe Ser Lys
                405                 410                 415
Gly Gly Gly Ser Pro Arg Ala Glu Ser Met Arg Pro Leu Thr Ser Tyr
            420                 425                 430
Tyr Gly Arg Gly Lys Arg Glu Ser Cys Val Ser Ser Ala Thr Thr Thr
        435                 440                 445
Thr Leu Val Gly Gln His Ala Pro Leu Ser Leu His Arg Gly Ser Asn
    450                 455                 460
Asn Ala Leu His Thr Met Pro Thr Leu Ala Ala Asn Ala Met Ser Ser
465                 470                 475                 480
Gly Ser Thr Leu Ser Val Met Pro Arg Ala Ile Ser Pro Leu Met Arg
                485                 490                 495
Gln Gly Leu Glu Glu Asn Ser Val
            500
```

<210> 103
<211> 3731
<212> DNA
<213> Drosophila

<400> 103

```
catcgattaa ccactatccc atttattgcc tgcccgtcta cacaaatcgg cacgtgaacg 60
aggactggta tcgactttat aggacctacg acaccgaggg atttgttttc gggcagtttt 120
acgagcgtct ccagcggtac gagttggact gaattaatgc ccttttttcag aaagaataag 180
atttagacaa aatctgtgat atcgatactg tggtagcgtt gcagagatcc caatcgcgtg 240
tgtggcgaag gacgcagaca gataacacat tcctccgtgc tgaaaatcaa ttataattag 300
tcgtgggaaa caactaaaag aatataaaaa tggcccgttg gtttataaaa aatataatac 360
tttaatctga tctcttataa tgtaaactat catgcaatac ttagggacga acggatgtgg 420
tctacatgta tgtcgtatac gtatattttt gtgtggtgtt cattcatagt gtccagtggg 480
tataggactt gttgatgtca gtctacaata aacgataaac agaaaaagga aaaccccaat 540
cagaaacata agaaaataag taagaataaa aaggatatct aaagccaacc aaccatgcag 600
aataaacata acattaagaa ctgatgtagg gaaatgcacg gttttaagta gatcacttta 660
gttggcacga gtgtgttagt atatttagag aacgggtgca atttaacttt ggaattattt 720
ggatgggtgg atgtcaaaga agtagaaagc atatacatga gcattgcgta tacgcaccgt 780
agactttcgg ttcgcgatct cagcgagtag tcttttgagg cgcgatgcct gaaatggaaa 840
ttaattggtt aaatgacaac aacaacaaca acaacatcga cgacaatagc gaacaactac 900
aatatcagcc atagccacat caactaggta gtgcaaatag tgaaaagcca atatgtcgac 960
agccagggtt atataaatat aatccaacaa caataatgtg ctagaatgcc aaatgagtgc 1020
agtcgcccct caatgattgg ttagcccacc tgcgattctt gatggaactg ttccgatgga 1080
tattcctgct ctcccgccac cgggtgaatc catgcctcag cgtctggcgc acctctgagt 1140
taaggaaaca gtagaacaga gccacaaaga agccctaaag aatggatttt cgaagatcaa 1200
tgtttatgga aagactgtca cctaatctat atacctgcgt gcttatgaga aaggctctta 1260
tggcctcgaa gagattacga ctgataccct gttccgggcc tgttagcacc aacagatagg 1320
tgatgccaaa gagaggtatc agcaccagca gcgccttcga ggccttgtaa tactgccgcg 1380
tttccagggt atgagcagaa cgtaatttag tgatgagtac ctgtacgata tgatataata 1440
```

gcatgaagct aactcactta agccagaagt taaccggggg actcacccac atgatgcgta 1500
tgaggaatac tagattaacc agcaaagcca gtgatgcagg taccttgaat atccagtcaa 1560
tgtgagattc acgcatccat gcacaatcaa tttccagctt ttgaaaaatc agaagaggtt 1620
agcaagcata tctaggtagg tggcaattgg tatcataccc cattgaagtg ctcgttctcg 1680
agatgcggag caaatgcctt ggcaatggac cacaccaaaa tgcatacggc tggacagccc 1740
cagccgatga gggcgtagat aataaagcta atgttatcac tggagaatgt ttgcaccacc 1800
agcgtgtaca gatagaggcc tataaaaagc aaatgataaa accgaggaaa ttcaattgat 1860
accgcctttg aacccaccct ccacaaacat ccagaaaaag ttggttaggt aaaagtactg 1920
aaacatgatt accaacgtta tgcagccagc ctgactagac tctgtggtta tctaagtgaa 1980
taattaatta gtggagtggg atatatatat tccagcttat cttaccactt gcaggaacag 2040
tgtgagtatc cagaggagtg cggatgtgat gtaggtgagg aacaaattgg catgaatggt 2100
gtttcgcagg caacgaagat ctctgaagag ttttaagcat gaaaatgatg caactcatgt 2160
gtggcattac aactcactta aagctgagga atatgatgag agccaccacc aaggtggcga 2220
agctcaggaa ataaccgccg gcatatatga tggccggcag ttcgacgttg ggtgagaagt 2280
cgggcaccac cggtatggag cccgagttct ggtgacagcg gtcatagtcc gaatagtgat 2340
cccacgttcc gtttggaaag caaaagcggg tggcattgtc tgcaaatata ttgattttat 2400
ttaagatcta aacccaatct ggaataccca cctgtggtgt cgtagtgcac gcccttgaat 2460
tcctcgaaac agggtagtac agccagactg ccggcgtttg tgcgtggcca gcacaggacc 2520
gaatcgaacg agctgggaca ttgtaaaatc tgtaagaaat aacaattgag aatatacaat 2580
gattggcaag gtaaacttcg ttgagcattt actgcagtat aattcgtgtt tcccctctgt 2640
aaacgcaact gacctagccc aaaacccatt aaatcagagt acataagatt catgccccgt 2700
ttctgactcg gtataataat gatgccttca ttgcggatcc tggtcattta tagctctgac 2760
gcatgatttt ctttaatcaa atttagtttc cagactttac acacacacat acgcagcgtg 2820
agccaaatat cgatgtagtc ccaaaaaaaa aataaagaaa gaatcaccag ctaccgcttt 2880
atttatatat ttgcaatggc gtgctaaaca atttcgtgca tacggataca caagttatat 2940
gtacatatat gtatattaac gttaagtata cgacgacccg gcgattgttt acccagttca 3000
gttcgtagcg ccaaattaat cgcgacgcgc tatggtagcg gtgttagaaa cagagcttca 3060
ttgatatttt ttttattgac cttatcaggt gacgccgaga gcgcggagcg gagactcgtc 3120
gaagaattca ccaaattggc aattctttta tggccggtct tcccattact gataagcccg 3180
agaacctttt tgcgacagct aatgcgatcg gacttagcca cgcgacaagt gagctgtcaa 3240
cgagtaagcc aaacactgga gactggaaac tccggcaata tttgtttctt tgctcttctg 3300
attttcactc attagcctcg tcatttggac acagtcatag tacgaataca ttttatgaca 3360
ttttcgagga acaacgccct tgatcgtgca ttcggtataa acaaaaatgg ctaaatgcag 3420
cctcaagtgg atttccttgg gtctagtttt gtttacaggc gagaggaata cataaataaa 3480
tatgtgctgc attttccgga aataaaatgc caaaataatt taatccgaaa gcgatcgaca 3540
gtttacttgc ctggtttttc atcgtatcgt acttaaagtg aaagagattg acgggcggat 3600
ggattgaagc caattatttc ggaacagaag atggatcaat tagcgggtgc gattaattgg 3660
gaaattcctc aatcggcaaa agagatgaat tcattggcca aacaaagggt aattagggaa 3720
tatttattgg c 3731

<210> 104
<211> 1167
<212> DNA
<213> Drosophila

<400> 104
atgaaggcat cattattata ccgaatttta caatgtccca gctcgttcga ttcggtcctg 60
tgctggccac gcacaaacgc cggcagtctg gctgtactac cctgtttcga ggaattcaag 120
ggcgtgcact acgacaccac agacaatgcc acccgctttt gctttccaaa cggaacgtgg 180
gatcactatt cggactatga ccgctgtcac cagaactcgg gctccatacc ggtggtgccc 240
gacttctcac ccaacgtcga actgccggcc atcatatatg ccggcggtta tttcctgagc 300
ttcgccacct tggtggtggc tctcatcata ttcctcagct ttaaagatct tcgttgcctg 360
cgaaacacca ttcatgccaa tttgttcctc acctacatca catccgcact cctctggata 420
ctcacactgt tcctgcaagt gataaccaca gagtctagtc aggctggctg cataacgttg 480
gtaatcatgt ttcagtactt ttacctaacc aactttttct ggatgtttgt ggagggcctc 540
tatctgtaca cgctggtggt gcaaacattc tccagtgata acattagctt tattatctac 600
gccctcatcg gctggggctg tccagccgta tgcattttgg tgtggtccat tgccaaggca 660
tttgctccgc atctcgagaa cgagcacttc aatgggctgg aaattgattg tgcatggatg 720
cgtgaatctc acattgactg gatattcaag gtacctgcat cactggcttt gctggttaat 780
ctagtattcc tcatacgcat catgtgggta ctcatcacta aattacgttc tgctcatacc 840

```
ctggaaacgc ggcagtatta caaggcctcg aaggcgctgc tggtgctgat acctctcttt 900
ggcatcacct atctgttggt gctaacaggc ccggaacagg gtatcagtcg taatctcttc 960
gaggccataa gagcctttct cataagcacg cagggcttct ttgtggctct gttctactgt 1020
ttccttaact cagaggtgcg ccagacgctg aggcatggat tcacccggtg gcgggagagc 1080
aggaatatcc atcggaacag ttccatcaag aatcgcaggt gggctaacca atcattgagg 1140
ggcgactgca ctcatttggc attctag                                     1167

<210> 105
<211> 388
<212> PRT
<213> Drosophila

<400> 105
Met Lys Ala Ser Leu Leu Tyr Arg Ile Leu Gln Cys Pro Ser Ser Phe
 1               5                  10                  15
Asp Ser Val Leu Cys Trp Pro Arg Thr Asn Ala Gly Ser Leu Ala Val
            20                  25                  30
Leu Pro Cys Phe Glu Glu Phe Lys Gly Val His Tyr Asp Thr Thr Asp
        35                  40                  45
Asn Ala Thr Arg Phe Cys Phe Pro Asn Gly Thr Trp Asp His Tyr Ser
    50                  55                  60
Asp Tyr Asp Arg Cys His Gln Asn Ser Gly Ser Ile Pro Val Val Pro
65                  70                  75                  80
Asp Phe Ser Pro Asn Val Glu Leu Pro Ala Ile Ile Tyr Ala Gly Gly
                85                  90                  95
Tyr Phe Leu Ser Phe Ala Thr Leu Val Val Ala Leu Ile Ile Phe Leu
            100                 105                 110
Ser Phe Lys Asp Leu Arg Cys Leu Arg Asn Thr Ile His Ala Asn Leu
        115                 120                 125
Phe Leu Thr Tyr Ile Thr Ser Ala Leu Leu Trp Ile Leu Thr Leu Phe
    130                 135                 140
Leu Gln Val Ile Thr Thr Glu Ser Ser Gln Ala Gly Cys Ile Thr Leu
145                 150                 155                 160
Val Ile Met Phe Gln Tyr Phe Tyr Leu Thr Asn Phe Phe Trp Met Phe
                165                 170                 175
Val Glu Gly Leu Tyr Leu Tyr Thr Leu Val Val Gln Thr Phe Ser Ser
            180                 185                 190
Asp Asn Ile Ser Phe Ile Ile Tyr Ala Leu Ile Gly Trp Gly Cys Pro
        195                 200                 205
Ala Val Cys Ile Leu Val Trp Ser Ile Ala Lys Ala Phe Ala Pro His
    210                 215                 220
Leu Glu Asn Glu His Phe Asn Gly Leu Glu Ile Asp Cys Ala Trp Met
225                 230                 235                 240
Arg Glu Ser His Ile Asp Trp Ile Phe Lys Val Pro Ala Ser Leu Ala
                245                 250                 255
Leu Leu Val Asn Leu Val Phe Leu Ile Arg Ile Met Trp Val Leu Ile
            260                 265                 270
Thr Lys Leu Arg Ser Ala His Thr Leu Glu Thr Arg Gln Tyr Tyr Lys
        275                 280                 285
Ala Ser Lys Ala Leu Leu Val Leu Ile Pro Leu Phe Gly Ile Thr Tyr
    290                 295                 300
Leu Leu Val Leu Thr Gly Pro Glu Gln Gly Ile Ser Arg Asn Leu Phe
305                 310                 315                 320
Glu Ala Ile Arg Ala Phe Leu Ile Ser Thr Gln Gly Phe Phe Val Ala
                325                 330                 335
Leu Phe Tyr Cys Phe Leu Asn Ser Glu Val Arg Gln Thr Leu Arg His
            340                 345                 350
Gly Phe Thr Arg Trp Arg Glu Ser Arg Asn Ile His Arg Asn Ser Ser
        355                 360                 365
Ile Lys Asn Arg Arg Trp Ala Asn Gln Ser Leu Arg Gly Asp Cys Thr
```

370 375 380
His Leu Ala Phe
385

<210> 106
<211> 7564
<212> DNA
<213> Drosophila

<400> 106
tttccgctgc catctgcagc tgtagctgcc taacaaatgc aattaaaatc tttgcgcaag 60
aaaattgttg cagcaggcac gcagagatac tcgtactgta tgtgctgcaa tagtttggtg 120
gcagcaacgg ggcggcagat aacatattgc ccacgtccgg ctgaaattgt acggacaagg 180
cgaactcaaa aaaaaaaaaa aaaaaaagga atccctcaaa aggagaacca agaaaaatgg 240
gatttatctc tggccaagct ccagccatac aaatctaatt tagaccacac gacggcttca 300
gatattttat attccccact cagttgcggt ccatcctgat ctcctccagc ttgcttgatg 360
agcccagccg ttggccaaat gaaattattt atgggcagtc ctaggtcagt gagaatttgg 420
ttttttaagg gtttgaactt tgcggctcaa caatgcctgt gatttttttt tttttatta 480
tttttttgca ttgcatccat tagagaaagc actcctttgg cattctactc agctacttgc 540
cacataaata tctacggagg gtagctaact agtccttatc tcaacttccc catttatcct 600
cgccgcaagt gcacttttcc tgctgctaat taagttgcaa ggcagctttc tcagcaggcg 660
ttgcacgaag tagataaaag gaaaagcccc cgatttgcag acactatgct gggaaagata 720
tatatatata tatattttgc cagcattcgc taatcggctt agggcagcca ccatgcgacc 780
acaacaacaa caacaacagc agcagcagtg tgctccatgc aatatttaat cacattagtt 840
gatggggtaa cggccatgta cgcaacaatt tacaacttgc attaattata cgtctagttg 900
gacaatggcg cctctaacta tcccactcta aatgtatctc tatctctatc cttatttgca 960
ggggacaccc gttgataata aattattaat gcatctgctg ggaaggcggc actgcacggc 1020
tgtccttggc agcaaatcca gttgcaacat cggtgagtga cgggcggcaa tcacaaagga 1080
cttgcaacag ttgccgttgc ttctgctatt gctgttgctg ttgctgttgc aacaatggcc 1140
agtggcaaca atgaaactga gccgctctac tgcggcagcg gcatggataa ttttcataca 1200
aggtaagcga ctttcaccaa tgtcacgatc aaggccaagg ggcttttcca ccagacatgg 1260
cgacaacaat accataagcc aacagttggg ggcgtggtcg ggggcggagg cgatggtggc 1320
aacatcatag ccattagctg ctggcgcaag ggaaccgttc aaaaatcgat tatcgcccca 1380
tttcgggggga gcttctattt tgatttgccg tacaattttc tcgggcgatt aaacgacgaa 1440
gcagaacgaa aacaaaaaac agatttgtca acagcaaggt caacaattga tggctgaaat 1500
caatttaatt gaccatatcc tacgggccct ccaagtggcc atctgctgca cctataaaaa 1560
agtgaatccg gtctgcgatt atttatatat tcgttgcatg gcaggcggtc gtaaaacctc 1620
gagatgatga ttaaaagcgg ccctaaaaac ttaatggcgg tttaggaaat tcaattcctg 1680
taatttaagc cgagtcacca ttcttcgaag ttcttacatg taagcgataa taaatagtta 1740
agtcaattgg ccaataaacc tattaatatt gtgcatttac cacgattaga ctttgattaa 1800
agtgacaatg ctgatttctg tagaggaaat ctagttctag tcttcccaca aagctattta 1860
gttactcttg aataaatatg ttacttttct tttgccaaaa ccaacagaat tttaaattta 1920
ataatttgga ttttttgcaa taaactgtac tgattaatgg gccacacaaa aatgtctagt 1980
ttattatgga gctcttggtt tcataaatta agaacataat ccaatcggca tataaatcat 2040
tgatagcaat ttattttccg tgatgaaact gtgctccgtg tgaactgcga attactcatt 2100
ctacggttgc aaaaaaagcc accaacggtc aacatttaga ccaggacttt tagttttaat 2160
tagagccagc ctggccaaca gcagtgttaa tgaccacaaa gtggctggcc acaggatcag 2220
catcccagaa tgcgatgccg catttgcttt aattaaaggt agtagctgga gtttgaaaga 2280
tgactgtatg gcaattagat gtgtagccag aacacttggc catttacttt tgtgtcaaag 2340
tcgtgccaaa ttgccagcgg aggcgacact tgacgctgtc acgccccaga cagacgcaga 2400
ccggcccaaa agcacccact cagccgtctc caggcgccac tcaagcggca aaggaacgcc 2460
aaaacactag gacacagaac gccagaagac tcgaaaaaaa agtataaaat tcaaacgcgt 2520
atacaattca atgtgtcaca tcgaagaaga accctttctc tggcggcaac gtgacgtata 2580
cgcaatatta atggcaaagc gaaaagggcc tcaaaagcta acgaaaaacc gcagcataac 2640
acgcacaaaa caagcaatta acataaataa taaagaacgc acaaaggcgt tcggctctcc 2700
gttgccaaag agaaaagcca acgaaaggaa taaaaaacgc caaaaagccc gaaaaaaaaa 2760
ccgaagaccc aaaaaatgtg taaaattgat ttcggaaaat atttaattta ccaaatgcat 2820
gctcgctttt tgcggccgtg tacgccccga acaaataaaa caaataggca cacatcgtct 2880
ggcatttcat gttaaaggct tctcatgctg atggacagga tacagaaaat atatatacga 2940

agtgaggtgt gtgtctacgt gccacaaaaa cagcgaaatt gaaacgcttt ccgaatgccg 3000
gcgacacttt ttggcggcac aaccctcgtc ttgtcgacca actcctgggg aatttacaac 3060
tccattcatc aaactcacaa aaactttgcc tggtgtacgg tcttgggttc gtcgtcgcag 3120
gatttcacct ggctaaaaat ccagactccc agccatccct gagtctaatt tatggccttt 3180
gataatccat gttgatgaat ggcggcaacg acaacaacgt cgacgacatg aatgaagtcc 3240
tggaattgtt ttgcaccagg atggcatcgg ggctccaggt gggacgtact ggctcaaagt 3300
tattggccca gaaatcaggc atagttagct gccgaaatga aacccaaata ccgagaaaac 3360
taggcaaaac aaacagtagt acaccggaaa tgcatatcat tgtaaaaact acatcagttt 3420
acctaaaagg cttggctttt aagctttcac atttataaaa tattgaaaat gcatataaaa 3480
gtatgaaatt aattcccttt tgtcaataaa ctttctttct ttctttctgt gtaatatggg 3540
ggataccggt tttttttttt tttcaatgaa atcccttcga aaggtataag ttcagaatcg 3600
agagttttat gccaagttgg gcacagtttt tttttcccc agctacctaa aataatagag 3660
acattttcct cccactacaa ctgattgcat tgccggtgca gaaagttttt tcagttggtt 3720
cggaaaaatt tggttcgcaa acaaattaat atgaactggc aagcattttt cgggcaaaaa 3780
gctctcatct atgtagattg gaatggaaat tccggctaga attgcataag accacctgca 3840
gtgtgggcta acatgactaa aaagttgtcc acaaatttgg cttagattct ccaataaaac 3900
tgtcgttcgg ccaggaatcc cctttttgt ttcgagtgaa tggggaattt cgcacgacag 3960
acagcaataa agaatttaac taaagtcctg acaccgacag caccagcagg acgcacacgt 4020
gtcactccat ttggagagct tggagtatat taaacatttt ttccccacca gtcagccgca 4080
ggacttgcat cggtctcgcc tcgcattttc ctatataaat tttatgctaa gtctaatttg 4140
ttggctgcaa cttgcacaaa ggcaaaaaat aaacaagggc gaaatgccga aagccaaaac 4200
ccaaccgaaa ccgttgaggg ctgcctcgct tttttcctgt gccgaattcc ctaaaacttt 4260
gcacataaat ttgagtcctg cgcctgggct tttcctcttc cacctttttt tttagcctca 4320
aagcgctcga cgaaaactaa acaaaagctg gtaaaatgtt tatcccagaa gggggcggtg 4380
tgggcgtggc ttgcttaagc cctagacacc cgagagttta ttgcctcagt gttgcaaacg 4440
ctgctgcttg gcaagtcaac ggaaatggct ttcggatggt tttattgcct acttttgctg 4500
cgagaaagcg gactagacat ggaagtaaca aattggttaa ggctaactag agcataataa 4560
cagaaacatc agaaacataa tttgggaaat gcctgccact caccgacaac atttcgattt 4620
gtgtctttat tacacttctc tgccgtgatg ggttataact tttggatcct tgcagagagt 4680
tgggaaattg cttaaaaata aaatttgttt aatttaattt gaggtggttt ttcaaacaat 4740
tttaatggat taataagttt aggaagttaa gtatatacat ctatctatta taccatcttc 4800
aaatattgtt gaattgaaga gataactcgg agctaatccc atatttttca tttttttcag 4860
ctacaagaac atgcatggct atgtttcgct ggtggtctgc atcctgggca ccatcgcgaa 4920
taccttgaat atcattgtgc taacccgacg ggagatgcgc tccccccacga atgccatact 4980
cacgggtctg gccgtggccg acctggcagt tatgctggag tatataccct acaccataca 5040
cgactacatc ctgacggaca gtttgccgcg ggaggagaag ctcagctaca gctgggcctg 5100
cttcatcaag ttccattcga ttttcgccca ggttctgcac accatttcca tttggctgac 5160
ggtgaccctg gctgtttggc gttatatagc ggtgggttat ccgcaaaaga atcgcgtatg 5220
gtgcggtatg agaaccacca taataacgat aaccaccgct tatgtggtgt gtgttctggt 5280
ggtgtcgccg tcgctctatt tgatcacggc tataaccgaa tatgtcgatc agttggatat 5340
gaatggcaaa gtgataaact ccattcccat gacccagtac gtaatcgatt atcgtaatga 5400
gttactgagt gccaggacgg ctgccctgaa tgccacgccc accagtgcac cactgaacga 5460
aactgtgtgg ttaaatgcga gcaccttgct gacatcgaca accaccgctg caccacccac 5520
gccatcgcca gtggtgcgaa atgttactgt ctataggcta taccacagcg atttggcgtt 5580
gcacaatgcc tcgctgcaaa atgccacatt tctcatatac agtgtagtga ttaagctgat 5640
accatgcata gcactcacca ttctgtcggt tcgattgatc ctggccttac tggaggccaa 5700
gcggcggcgg aagaagctca ccagcaagcc cgccactccg ggtgccagta atggaaccaa 5760
atcaccggcc aatggtaaag cagcggatag gccccggaaa aatagcaaaa ctctggaaaa 5820
ggaaaagcag acggatcgca ccacgagaat gctgctggcg gtgctactcc tctttctcat 5880
cactgaattt ccacaaggga ttatgggtct gctgaatgct gtgctcggag atgtcttcta 5940
tctgcagtgc tacctaagac tgagtaagta aatatgtact atgcataacg ccacagtaaa 6000
ttccacagat tcccatttca ctttaagatg taataaccca atgccttttg ccttttgcca 6060
ccaggtgacc tgatggatat cttggccctg atcaactcca gcatcaactt cattttgtac 6120
tgctccatga gcaagcaatt ccgcaccacg ttcacgctgc tctttcgtcc aaaattcctg 6180
gacaagtggc tgccggtggc gcaggacgaa atggcagctg ctcgagctga acgctctgcg 6240
gtggcaccgg tcctggaaaa gggacgacag cagccgcagg tggtgatggc cagcacgacc 6300
accaacatca cgcaggtgac aaatctgtag cacaggagga gtcgtggtcg ccgaactttg 6360
ctcagtcgcc tgctgagtgt cctgaaacgc ggcagaaggc gctcctccgg cgaaggagga 6420
ggcgtgggag gaggtggtgc gccgttggcc ggcaacgatg cggtggaacc agcgttccag 6480
gccatcgtgg tggtggtgga caaggtgagc ggtgccacgg agaatcagct gtacaccgcc 6540

gagcaagctc gtattgtgac gtagtaaccc tgatagtagt acatattata tgcctagtta 6600
tccttgtaaa aagtaaacta gttgtgtaag aatcgcagat aaccacaagt tatgtgcata 6660
ctttggtggg aaatcgtaaa acccgtacag tcataataat aaaatgctgg ttcatcactc 6720
atacgcaccg tttgacttca ctctattcca cattcacagt caagtaggag ctgtaaaaat 6780
gctgtgcata aacgagctcg tatacgctac agctggcgtc acaataatag ttaccattgg 6840
gttggcttca caacatagat caaagtcaaa aaatatatag ttttgtaaaa tgtttcataa 6900
atctatcttc atatttaact aacacaattc agaaactttc gcttgcatgg aaaatgtttt 6960
attttcccca cttgacccaa ttgtgcgcaa ctgttcagct tttcggcatg aatcaaatca 7020
aatatagtaa acttttgcat cttttccact ttgatcccat gcgttttcat gttgatttca 7080
ttcgcataaa gtatgcagca cacgtactca ctcaaaacta cgaattttcg gcaacgattt 7140
gagcagaata tttgccattt gttgtacaga ttatatatac acacacacac acttatgtcc 7200
gtatctttaa taatggcaat taagtcctag aaattttgac ctgatattga tgcaggcatt 7260
ggaaaaataa gtggcagttg cagtgcggtg cgttgataaa tgaaatactt tgcgaattct 7320
attattggaa atcgcaattt gggctgtgaa tgaaattata gttttcttga acgcaattga 7380
gaattcagtt ccgtttcgtg ggttactgaa tttatgggaa ataatttaag gcagtcactg 7440
cagctgaccg aataatccaa tgaaacattg cataaacagt ttgcagtttt tccacacttg 7500
cgaaaattgt aattgcaaat cgcttaagcc ccacgacgaa tgacaaatgg ccagcgaacg 7560
cata 7564

<210> 107
<211> 1757
<212> DNA
<213> Drosophila

<400> 107
ggaaggcggc actgcacggc tgtccttggc agcaaatcca gttgcaacat cggtgagtga 60
cgggcggcaa tcacaaagga cttgcaacag ttgccgttgc ttctgctatt gctgttgctg 120
ttgctgttgc aacaatggcc agtggcaaca atgaaactga gccgctctac tgcggcagcg 180
gcatggataa ttttcataca agctacaaga acatgcatgg ctatgtttcg ctggtggtct 240
gcatcctggg caccatcgcg aataccttga atatcattgt gctaacccga cgggagatgc 300
gctcccccac gaatgccata ctcacgggtc tggccgtggc cgacctggca gttatgctgg 360
agtatatacc ctacaccata cacgactaca tcctgacgga cagtttgccg cgggaggaga 420
agctcagcta cagctgggcc tgcttcatca agttccattc gattttcgcc caggttctgc 480
acaccatttc catttggctg acggtgaccc tggctgtttg gcgttatata gcggtgggtt 540
atccgcaaaa gaatcgcgta tggtgcggta tgagaaccac cataataacg ataaccaccg 600
cttatgtggt gtgtgttctg gtggtgtcgc cgtcgctcta tttgatcacg gctataaccg 660
aatatgtcga tcagttggat atgaatggca aagtgataaa ctccattccc atgacccagt 720
acgtaatcga ttatcgtaat gagttactga gtgccaggac ggctgccctg aatgccacgc 780
ccaccagtgc accactgaac gaaactgtgt ggttaaatgc gagcaccttg ctgacatcga 840
caaccaccgc tgcaccaccc acgccatcgc cagtggtgcg aaatgttact gtctataggc 900
tataccacag cgatttggcg ttgcacaatg cctcgctgca aaatgccaca tttctcatat 960
acagtgtagt gattaagctg ataccatgca tagcactcac cattctgtcg gttcgattga 1020
tcctggcctt actggaggcc aagcggcggc ggaagaagct caccagcaag cccgccactc 1080
cgggtgccag taatggaacc aaatcaccgg ccaatggtaa agcagcggat aggccccgga 1140
aaaatagcaa aactctggaa aaggaaaagc agacggatcg caccacgaga atgctgctgg 1200
cggtgctact cctctttctc atcactgaat ttccacaagg gattatgggt ctgctgaatg 1260
ctgtgctcgg agatgtcttc tatctgcagt gctacctaag actgagtgac ctgatggata 1320
tcttggccct gatcaactcc agcatcaact tcattttgta ctgctccatg agcaagcaat 1380
tccgcaccac gttcacgctg ctctttcgtc caaaattcct ggacaagtgg ctgccggtgg 1440
cgcaggacga aatggcagct gctcgagctg aacgctctgc ggtggcaccg gtcctggaaa 1500
agggacgaca gcagccgcag gtgcacagga ggagtcgtgg tcgccgaact ttgctcagtc 1560
gcctgctgag tgtcctgaaa cgcggcagaa ggcgctcctc cggcgaagga ggaggcgtgg 1620
gaggaggtgg tgcgccgttg gccggcaacg atgcggtgga accagcgttc caggccatcg 1680
tggtggtggt ggacaaggtg agcggtgcca cggagaatca gctgtacacc gccgagcaag 1740
ctcgtattgt gacgtag 1757

<210> 108
<211> 540

<212> PRT
<213> Drosophila

<400> 108
Met Ala Ser Gly Asn Asn Glu Thr Glu Pro Leu Tyr Cys Gly Ser Gly
1 5 10 15
Met Asp Asn Phe His Thr Ser Tyr Lys Asn Met His Gly Tyr Val Ser
20 25 30
Leu Val Val Cys Ile Leu Gly Thr Ile Ala Asn Thr Leu Asn Ile Ile
35 40 45
Val Leu Thr Arg Arg Glu Met Arg Ser Pro Thr Asn Ala Ile Leu Thr
50 55 60
Gly Leu Ala Val Ala Asp Leu Ala Val Met Leu Glu Tyr Ile Pro Tyr
65 70 75 80
Thr Ile His Asp Tyr Ile Leu Thr Asp Ser Leu Pro Arg Glu Glu Lys
85 90 95
Leu Ser Tyr Ser Trp Ala Cys Phe Ile Lys Phe His Ser Ile Phe Ala
100 105 110
Gln Val Leu His Thr Ile Ser Ile Trp Leu Thr Val Thr Leu Ala Val
115 120 125
Trp Arg Tyr Ile Ala Val Gly Tyr Pro Gln Lys Asn Arg Val Trp Cys
130 135 140
Gly Met Arg Thr Thr Ile Ile Thr Ile Thr Thr Ala Tyr Val Val Cys
145 150 155 160
Val Leu Val Val Ser Pro Ser Leu Tyr Leu Ile Thr Ala Ile Thr Glu
165 170 175
Tyr Val Asp Gln Leu Asp Met Asn Gly Lys Val Ile Asn Ser Ile Pro
180 185 190
Met Thr Gln Tyr Val Ile Asp Tyr Arg Asn Glu Leu Leu Ser Ala Arg
195 200 205
Thr Ala Ala Leu Asn Ala Thr Pro Thr Ser Ala Pro Leu Asn Glu Thr
210 215 220
Val Trp Leu Asn Ala Ser Thr Leu Leu Thr Ser Thr Thr Thr Ala Ala
225 230 235 240
Pro Pro Thr Pro Ser Pro Val Val Arg Asn Val Thr Val Tyr Arg Leu
245 250 255
Tyr His Ser Asp Leu Ala Leu His Asn Ala Ser Leu Gln Asn Ala Thr
260 265 270
Phe Leu Ile Tyr Ser Val Val Ile Lys Leu Ile Pro Cys Ile Ala Leu
275 280 285
Thr Ile Leu Ser Val Arg Leu Ile Leu Ala Leu Leu Glu Ala Lys Arg
290 295 300
Arg Arg Lys Lys Leu Thr Ser Lys Pro Ala Thr Pro Gly Ala Ser Asn
305 310 315 320
Gly Thr Lys Ser Pro Ala Asn Gly Lys Ala Ala Asp Arg Pro Arg Lys
325 330 335
Asn Ser Lys Thr Leu Glu Lys Glu Lys Gln Thr Asp Arg Thr Thr Arg
340 345 350
Met Leu Leu Ala Val Leu Leu Leu Phe Leu Ile Thr Glu Phe Pro Gln
355 360 365
Gly Ile Met Gly Leu Leu Asn Ala Val Leu Gly Asp Val Phe Tyr Leu
370 375 380
Gln Cys Tyr Leu Arg Leu Ser Asp Leu Met Asp Ile Leu Ala Leu Ile
385 390 395 400
Asn Ser Ser Ile Asn Phe Ile Leu Tyr Cys Ser Met Ser Lys Gln Phe
405 410 415
Arg Thr Thr Phe Thr Leu Leu Phe Arg Pro Lys Phe Leu Asp Lys Trp
420 425 430
Leu Pro Val Ala Gln Asp Glu Met Ala Ala Ala Arg Ala Glu Arg Ser
435 440 445

```
Ala Val Ala Pro Val Leu Glu Lys Gly Arg Gln Gln Pro Gln Val His
    450                 455                 460
Arg Arg Ser Arg Gly Arg Arg Thr Leu Leu Ser Arg Leu Leu Ser Val
465                 470                 475                 480
Leu Lys Arg Gly Arg Arg Arg Ser Ser Gly Glu Gly Gly Gly Val Gly
                485                 490                 495
Gly Gly Gly Ala Pro Leu Ala Gly Asn Asp Ala Val Glu Pro Ala Phe
            500                 505                 510
Gln Ala Ile Val Val Val Val Asp Lys Val Ser Gly Ala Thr Glu Asn
        515                 520                 525
Gln Leu Tyr Thr Ala Glu Gln Ala Arg Ile Val Thr
    530                 535                 540
```

```
<210> 109
<211> 5251
<212> DNA
<213> Drosophila

<400> 109
taaccttctt ttgaccttcg taagtcttcc actggcactt gaaaagcttt gggaaaacag 60
ctgtgaattt gcccagcaga gcttccgaaa cgacggcttt tagtacctcc gccgtaacac 120
cacccacagt gggtcggttc tggtccatgg tggggatggt gtggtggttc cgatgatacg 180
ggaatcgctg gagaagcgtt actgatgcaa tgcactggat accgttcctg ttcagatcat 240
tgagtactcc aggcggcatt tctattgtgg ttacaatagt aagtaaggag tagtaaaagt 300
tattctgaaa ttatataaac ggcttgttgg ggttttaaaa atgttgccgg ttttagattt 360
agcagtgact ccaaccatct gatggggaac tttgagcaag tgagctccaa tagttttggc 420
gtcgcttcca aaaagctctg gaaaagcttt tatcaacagg cttgccaaaa tccgctatac 480
tagcatatgt aaaagaaaaa cttgtatacc ctaccaatag gcatacatcg catcaggact 540
tatttaacca acaaaaataa aacacaacat gaatgtacat tctggtttaa tcgctttatt 600
taacttcaac tttgaaattc atgtcgcatg gcttaacttg gttttcacag acgttcaggt 660
atacataagc tgacatatac acagtaaata ctctaacaac attaaacata tgtatgtaca 720
taattaacta gtccgtagaa acgtattatg accatacact caaaataaat gcgcatagct 780
cctaatctct atatttcact accgaactaa ttgatttaga tccgatatag aatttgacaa 840
cgtttccgaa agagagccta agagagctgg aaagtactta acgtaaaagc tttgacttgc 900
gagtatttcc tcaggggtca cccccttaccc ctttcgcatc cttgtgatcc tcttccagtg 960
gactttcctg ctgatcccgc ttcttttccg accgcgaagg tcaactctcc gcgctggtcg 1020
ccagaaggga cggcgaggag gtggtcagac ccactggaac cgattcctcc tgagcctggc 1080
ccatgggcgg tggtgatgga gtgggcgtgg aggccgtgga cgacgccgtc gatgaggtgg 1140
gacgggcggc gtgaggctgg aggaagtgaa acaagttgga atgctgcatg aaactgccca 1200
ccgatctcag cggactaaaa aagttctctg tgggatggtt ggctggctgc ggatgctgtg 1260
gaagtacaga ggactcctga tccctgcaga gattactgct gggctcgtcg tggacccctc 1320
gtcccttagc gatggagcta cagttgcagg atcggctctt cttcagcgaa ctccgagtga 1380
atgccgtgcc ctttcgatac ttaggcttgt gcccatttga gccgggactc gcctcactat 1440
cgatcacaaa gataccggat tgctcctggg gagcaggact cgccgtggat cctggcaggc 1500
agtgagcgct attggagttg ccagcaagca gtttggccct cagaagagcc ggaagtgggc 1560
catttgtggg agatgtttga gtggagagcg gtctctccct ggactgttca tctatgggaa 1620
gatcctgatt gttgcccatt ggcctggagt ccgtgtcatc cgatcgcaat ttctccctca 1680
tgaccaaagc cagggagcga cgcagttcca cgggatctat gccaggctga acgggatttg 1740
gaacagagct ggccggcacg aatacgtcct caaactccag atcatcatca tcttcgtcct 1800
catcctggct gaaactgtag ttgtggttgt cataggaggc ggacggaact gagggcttcg 1860
caggagcagg aggagctggt gggactgtgg aactgctatt tcgttcactg ctgcgctggg 1920
aacccggagc gggggaaggg aactcgtagc tggtttcacc cagctcggcg ccactggagg 1980
cctccactat gttccacatt tctcgctcga gatctgtggt caccgaagcc acggcggcca 2040
ccgtttggtt gatgttgttg aaacagggcg acaccactcc gcacttgggc aggagctggg 2100
agttcttcag ggcctcctgc tgctgcctca accaaatggg tgaatctgca gggaatgtta 2160
cgagagttat agggattttg ccacagttcc aaggagacaa cgcgcactca cactcaccca 2220
catagaggcg cggctgcgtg gagcgcatat tggagtccat ggactgcagg cgatgctgac 2280
tggccaccga gaagcggtgg attgcctccg cctgctgatc aatgtccacg cccatcatct 2340
tcaggagtaa gttcttcagc gcggccctga agtccttgag gtgataggca tacaggaccg 2400
```

gatttaccgc cgagttcaga tgggacagaa tgatgcagaa gagcgtcagc ttgggatgca 2460
catagcagtc gggacagaag gccttgatgc agttgatcgt atagagcggt atccagcaga 2520
tcataaagaa cagcacgata atggacagat tctgggtggc cttgacgtcc cgcttccttg 2580
cggcgcccaa gacccgcagc atggttcccg tgtgtccacc tcttccgggt gtcgtcacct 2640
gcaccacggc cgccgaggag cgacgactga gatcggaggc ggggttcatc gtaacaatct 2700
gacggacctt taaagttcaa ttgaaaacgg atagttaata ttccaaactc tatggaatat 2760
ataaaaaact ttactggcga aaatctatct accactcacc tgtttgatga tgacccggta 2820
gatgtgcgtg tagaaggcca gcatcagcag agccggagtg ataatggtgg caaaatagag 2880
gaagacgagg tagttgtagt ccatcacctc cacgaagaga cactcctggt tgtgattgac 2940
atcggcgtgc cagccgaaaa gcggcaggaa gcccactatc gtcccggcca cccagcacat 3000
ggagatgatg actgcaacgg aagaggaaca ttggtaatca gagtgccctg aactgaagcc 3060
cttgaagtgg agctccccca gacgtcattt gagcagattg caggaatccc tgccgtcttc 3120
gccaccaagc accataaaaa cttttatcga attgcgggaa actaaacaaa actggcaggc 3180
agggtgagtg cacttcccaa atgacgcaaa taattaccaa ttttccggac gtagcatatt 3240
ggaaagtaat cttgccttga gcggccattt ataaagcact atatagtagg tgatatatag 3300
tgcttagaaa attatatatt tatgagatat ttttagctag aggccaatta ttctgattcc 3360
aaacagttgc tataaacatc taatatgaaa attcttagaa tcaatttgaa tctagttagg 3420
ctagtaagac tattatgtta cggattttaa ggatgtttat tactgtgtac tgttctcatt 3480
gcgtccattc gtggtagttc aaatgaggcg tcttttgttt aactcaagtc gcaactcaaa 3540
catggtcgat ggccggaatg aaacacaatg tgctgtatat tagccatcca gagtcattag 3600
tatcacggca acaaaggcac aatactaaac aatggacaac ttaaagtcga aatagtttga 3660
caaatatcct gacaagtagg tgaaaccaga aaactggtcc tttaaagtgg gttttagcat 3720
accattttgt acctactgat tttatacaat acattgtata tacccccata tatacttaca 3780
tatcgccgtg cgggtgcgga cattccttga gtaggccatc ggatatagga tggcccagta 3840
tcgatccacg gacacggcca ccagacagaa gatggagatg gtacacagca ccacgagcag 3900
ggagactgtg aagaggcagg catgaaggtt tctgggtagt cccatggagg ccaggatggc 3960
gaagggaata cccaatgcgc ccactagcaa atcggccatg gccagggata ctatgtagta 4020
gttggtgcgg cgtcgcaact ttctttcccg gcggaagacg atgatcacca gcacgttgcc 4080
gatgatgctg acgatggcca ccaggacctc gaagacggtg tagggtatgt tcagctcgga 4140
actggggctg tccgagtcct tggcatcctt gctcgtggtg gccgcgtgca gcggtagcag 4200
aggtccctcg aaggagaaat cggtgatcga gaagtagcga aacgcggaca tggtgggcgg 4260
tgttgcgggt attttacggg ttccacgggg ttccctgccc cgatcgagta tttaaggcta 4320
gtgccgagct aatgaaaggg cggaacaggt gtgggcttca attacagtgc ccgctggcct 4380
gtgaactccc taattgactt gaggtggccg tggcagtagc agcgatagta atagcacatg 4440
ctgcgcattg cgtatacgcc atgcccgccg cgctgcggtt gctcctttcc cgctttctcg 4500
cttttccgat tcgctttttt cgctctcgct ctttatttc ctttccgttg atgtatttgt 4560
ttttctgctg tcgacgtacc ctattaagct ggaaaaaaaa gggaaaactt aatcatcgtt 4620
cggcacgttt gttacacaca tgcagcaact taattgtagc aattttttag tgtaataaat 4680
tgataaatat ttgccaaaaa atacatgtcc atattgataa cgtgtgcggc gagcaaaaca 4740
acaaaggccg tatcacaact atgcaatacc ctagataatc ggccagtact tgagtatatt 4800
atttgaaagg ggccatcagt tgcaggtctt ccttaaagcc agcagctcaa ggatcacagt 4860
gaaatatgtt ggaacttcga tggcttttac gatgatggtg tttctggttt ggcttatggg 4920
gtatatatac ccgcaactct gcgagtacac actgccacac tatctgattc gcgtgcggca 4980
aatacaaagc tcgaaaagca aacaactccc gactagcaac aaaatcaatt gccagacaat 5040
cgaaccgggc gctgcagaac ataataaaag tgtataaaaa ggacggaaca gcgagaacaa 5100
cgcagaaaca gacgtcccat aaataaaaga gcatcctaca cagagaaaaa agaattaggc 5160
aggggtttac gaaaataatt catattcaat tcaaaagtat tcgctttttt tgtaattatt 5220
aagatagatg atgtattgct cctaaaaact c 5251

<210> 110
<211> 2325
<212> DNA
<213> Drosophila

<400> 110
atgtccgcgt ttcgctactt ctcgatcacc gatttctcct tcgagggacc tctgctaccg 60
ctgcacgcgg ccaccacgag caaggatgcc aaggactcgg acagccccag ttccgagctg 120
aacataccct acaccgtctt cgaggtcctg gtggccatcg tcagcatcat cggcaacgtg 180

ctggtgatca tcgtcttccg ccgggaaaga aagttgcgac gccgcaccaa ctactacata 240
gtatccctgg ccatggccga tttgctagtg ggcgcattgg gtattccctt cgccatcctg 300
gcctccatgg gactacccag aaaccttcat gcctgcctct tcacagtctc cctgctcgtg 360
gtgctgtgta ccatctccat cttctgtctg gtggccgtgt ccgtggatcg atactggcc 420
atcctatatc cgatggccta ctcaaggaat gtccgcaccc gcacggcgat attcatcatc 480
tccatgtgct gggtggccgg gacgatagtg ggcttcctgc cgcttttcgg ctggcacgcc 540
gatgtcaatc acaaccagga gtgtctcttc gtggaggtga tggactacaa ctacctcgtc 600
ttcctctatt ttgccaccat tatcactccg gctctgctga tgctggcctt ctacacgcac 660
atctaccggg tcatcatcaa acaggtccgt cagattgtta cgatgaaccc cgcctccgat 720
ctcagtcgtc gctcctcggc ggccgtggtg caggtgacga cacccggaag aggtggacac 780
acgggaacca tgctgcgggt cttgggcgcc gcaaggaagc gggacgtcaa ggccacccag 840
aatctgtcca ttatcgtgct gttctttatg atctgctgga taccgctcta tacgatcaac 900
tgcatcaagg ccttctgtcc cgactgctat gtgcatccca agctgacgct cttctgcatc 960
attctgtccc atctgaactc ggcggtaaat ccggtcctgt atgcctatca cctcaaggac 1020
ttcagggccg cgctgaagaa cttactcctg aagatgatgg gcgtggacat tgatcagcag 1080
gcggaggcaa tccaccgctt ctcggtggcc agtcagcatc gcctgcagtc catggactcc 1140
aatatgcgct ccacgcagcc gcgcctctat gtgggtgagt attcacccat ttggttgagg 1200
cagcagcagg aggccctgaa gaactcccag ctcctgccca agtgcggagt ggtgtcgccc 1260
tgtttcaaca acatcaacca aacggtggcc gccgtggctt cggtgaccac agatctcgag 1320
cgagaaatgt ggaacatagt ggaggcctcc agtggcgccg agctgggtga aaccagctac 1380
gagttccctt cccccgctcc gggttcccag cgcagcagtg aacgaaatag cagttccaca 1440
gtcccaccag ctcctcctgc tcctgcgaag ccctcagttc cgtccgcctc ctatgacaac 1500
cacaactaca gtttcagcca ggatgaggac gaagatgatg atgatctgga gtttgaggac 1560
gtattcgtgc cggccagctc tgttccaaat cccgttcagc ctggcataga tcccgtggaa 1620
ctgcgtcgct ccctggcttt ggtcatgagg gagaaattgc gatcggatga cacggactcc 1680
aggccaatgg gcaacaatca ggatcttccc atagatgaac agtccaggga gagaccgctc 1740
tccactcaaa catctcccac aaatggccca cttccggctc ttctgagggc caaactgctt 1800
gctggcaact ccaatagcgc tcactgcctg ccaggatcca cggcgagtcc tgctccccag 1860
gagcaatccg gtatctttgt gatcgatagt gaggcgagtc ccggctcaaa tgggcacaag 1920
cctaagtatc gaaagggcac ggcattcact cggagttcgc tgaagaagag ccgatcctgc 1980
aactgtagct ccatcgctaa gggacgaggg gtccacgacg agcccagcag taatctctgc 2040
agggatcagg agtcctctgt acttccacag catccgcagc cagccaacca tcccacagag 2100
aactttttta gtccgctgag atcggtgggc agtttcatgc agcattccaa cttgtttcac 2160
ttcctccagc ctcacgccgc ccgtcccacc tcatcgacgg cgtcgtccac ggcctccacg 2220
cccactccat caccaccgcc catgggccag gctcaggagg aatcggttcc agtgggtctg 2280
accacctcct cgccgtccct tctggcgacc agcgcggaga gttga 2325

<210> 111
<211> 774
<212> PRT
<213> Drosophila

<400> 111
Met Ser Ala Phe Arg Tyr Phe Ser Ile Thr Asp Phe Ser Phe Glu Gly
1 5 10 15
Pro Leu Leu Pro Leu His Ala Ala Thr Thr Ser Lys Asp Ala Lys Asp
20 25 30
Ser Asp Ser Pro Ser Ser Glu Leu Asn Ile Pro Tyr Thr Val Phe Glu
35 40 45
Val Leu Val Ala Ile Val Ser Ile Ile Gly Asn Val Leu Val Ile Ile
50 55 60
Val Phe Arg Arg Glu Arg Lys Leu Arg Arg Arg Thr Asn Tyr Tyr Ile
65 70 75 80
Val Ser Leu Ala Met Ala Asp Leu Leu Val Gly Ala Leu Gly Ile Pro
85 90 95
Phe Ala Ile Leu Ala Ser Met Gly Leu Pro Arg Asn Leu His Ala Cys
100 105 110
Leu Phe Thr Val Ser Leu Leu Val Val Leu Cys Thr Ile Ser Ile Phe
115 120 125
Cys Leu Val Ala Val Ser Val Asp Arg Tyr Trp Ala Ile Leu Tyr Pro

```
    130                 135                 140
Met Ala Tyr Ser Arg Asn Val Arg Thr Arg Thr Ala Ile Phe Ile Ile
145                 150                 155                 160
Ser Met Cys Trp Val Ala Gly Thr Ile Val Gly Phe Leu Pro Leu Phe
                165                 170                 175
Gly Trp His Ala Asp Val Asn His Asn Gln Glu Cys Leu Phe Val Glu
            180                 185                 190
Val Met Asp Tyr Asn Tyr Leu Val Phe Leu Tyr Phe Ala Thr Ile Ile
        195                 200                 205
Thr Pro Ala Leu Leu Met Leu Ala Phe Tyr Thr His Ile Tyr Arg Val
    210                 215                 220
Ile Ile Lys Gln Val Arg Gln Ile Val Thr Met Asn Pro Ala Ser Asp
225                 230                 235                 240
Leu Ser Arg Arg Ser Ser Ala Ala Val Val Gln Val Thr Thr Pro Gly
                245                 250                 255
Arg Gly Gly His Thr Gly Thr Met Leu Arg Val Leu Gly Ala Ala Arg
            260                 265                 270
Lys Arg Asp Val Lys Ala Thr Gln Asn Leu Ser Ile Ile Val Leu Phe
        275                 280                 285
Phe Met Ile Cys Trp Ile Pro Leu Tyr Thr Ile Asn Cys Ile Lys Ala
    290                 295                 300
Phe Cys Pro Asp Cys Tyr Val His Pro Lys Leu Thr Leu Phe Cys Ile
305                 310                 315                 320
Ile Leu Ser His Leu Asn Ser Ala Val Asn Pro Val Leu Tyr Ala Tyr
                325                 330                 335
His Leu Lys Asp Phe Arg Ala Ala Leu Lys Asn Leu Leu Leu Lys Met
            340                 345                 350
Met Gly Val Asp Ile Asp Gln Gln Ala Glu Ala Ile His Arg Phe Ser
        355                 360                 365
Val Ala Ser Gln His Arg Leu Gln Ser Met Asp Ser Asn Met Arg Ser
    370                 375                 380
Thr Gln Pro Arg Leu Tyr Val Gly Glu Tyr Ser Pro Ile Trp Leu Arg
385                 390                 395                 400
Gln Gln Gln Glu Ala Leu Lys Asn Ser Gln Leu Leu Pro Lys Cys Gly
                405                 410                 415
Val Val Ser Pro Cys Phe Asn Asn Ile Asn Gln Thr Val Ala Ala Val
            420                 425                 430
Ala Ser Val Thr Thr Asp Leu Glu Arg Glu Met Trp Asn Ile Val Glu
        435                 440                 445
Ala Ser Ser Gly Ala Glu Leu Gly Glu Thr Ser Tyr Glu Phe Pro Ser
    450                 455                 460
Pro Ala Pro Gly Ser Gln Arg Ser Ser Glu Arg Asn Ser Ser Ser Thr
465                 470                 475                 480
Val Pro Pro Ala Pro Pro Ala Pro Ala Lys Pro Ser Val Pro Ser Ala
                485                 490                 495
Ser Tyr Asp Asn His Asn Tyr Ser Phe Ser Gln Asp Glu Asp Glu Asp
            500                 505                 510
Asp Asp Asp Leu Glu Phe Glu Asp Val Phe Val Pro Ala Ser Ser Val
        515                 520                 525
Pro Asn Pro Val Gln Pro Gly Ile Asp Pro Val Glu Leu Arg Arg Ser
    530                 535                 540
Leu Ala Leu Val Met Arg Glu Lys Leu Arg Ser Asp Asp Thr Asp Ser
545                 550                 555                 560
Arg Pro Met Gly Asn Asn Gln Asp Leu Pro Ile Asp Glu Gln Ser Arg
                565                 570                 575
Glu Arg Pro Leu Ser Thr Gln Thr Ser Pro Thr Asn Gly Pro Leu Pro
            580                 585                 590
Ala Leu Leu Arg Ala Lys Leu Leu Ala Gly Asn Ser Asn Ser Ala His
        595                 600                 605
Cys Leu Pro Gly Ser Thr Ala Ser Pro Ala Pro Gln Glu Gln Ser Gly
```

```
    610                 615                 620
Ile Phe Val Ile Asp Ser Glu Ala Ser Pro Gly Ser Asn Gly His Lys
625                 630                 635                 640
Pro Lys Tyr Arg Lys Gly Thr Ala Phe Thr Arg Ser Ser Leu Lys Lys
                645                 650                 655
Ser Arg Ser Cys Asn Cys Ser Ser Ile Ala Lys Gly Arg Gly Val His
            660                 665                 670
Asp Glu Pro Ser Ser Asn Leu Cys Arg Asp Gln Glu Ser Ser Val Leu
        675                 680                 685
Pro Gln His Pro Gln Pro Ala Asn His Pro Thr Glu Asn Phe Phe Ser
    690                 695                 700
Pro Leu Arg Ser Val Gly Ser Phe Met Gln His Ser Asn Leu Phe His
705                 710                 715                 720
Phe Leu Gln Pro His Ala Ala Arg Pro Thr Ser Ser Thr Ala Ser Ser
                725                 730                 735
Thr Ala Ser Thr Pro Thr Pro Ser Pro Pro Pro Met Gly Gln Ala Gln
            740                 745                 750
Glu Glu Ser Val Pro Val Gly Leu Thr Thr Ser Ser Pro Ser Leu Leu
        755                 760                 765
Ala Thr Ser Ala Glu Ser
    770
```

<210> 112
<211> 4314
<212> DNA
<213> Drosophila

<400> 112

```
gagccaattc tttttggct ttacgtgacg ttggcagaac aaatcgtttt tgttttagcc 60
atctgacttg tttgctcctt tatggcccgc tgcgctttgg actttgacgc atttgcgtcg 120
ctttaatgcg attgctggtt gctcattacg gccaatcctt actagtactt gcactttctg 180
tgaattcgtc tggggtttcc atttggctga ttgcattttc ataacaattt ggaggattac 240
taatgcgcca cgttgtccct gcatttgcat gaccttgaac cagtagccac ctcattagaa 300
tttcatttgt agtcaacaag gcgcagacga cagggccaca atcgataaaa acccaaatgt 360
attttatggc caaattacat ttttactaga atgaatccat agaaagatgc ggaaaataag 420
cgggcgcgga atatgctaag ctttatttgc agccatggga ataatcttaa tacagcaaat 480
tcatccttta aggtatctct agaagtaagt ttagtaagcg aactttgtcc gtcatttcca 540
ttgagttgaa tgctacattt tcgaaatctt attttaaatg cgattcaata caatatacag 600
aatctactat gtttaagatc gcaacatttt tttttacact ttcttaacta aacactttaa 660
aagcacatct gtgttacact ttggtatgca tgtgaaacat ttggttaact aaaacatata 720
tctacgcaat ttaaaaggcg atctattaaa tttcaagggg caagcttttc gtttttaaat 780
ttaattccta tatttaaatg gtgggcgttc actagtgatt gatttatgct taccattatc 840
gataatatta ttaagttata tccgcctcag gcttgtttaa gagtaaatcc cccgagctct 900
gaccaaagtt taattttaag agcgccagta tgctgaatag tatgagtcaa aagattgaat 960
attcttagac tgaataattc tgaatacttg aaagactagt ttaggagttt ttccagccgt 1020
tgtccgcagc tgaatgaatg aaaaacttaa tgcttaaggg agagtgtgcg ctcgactgtg 1080
attaactttt aattgccaaa tgcaaatgaa tgcaaaagca gtttaaacaa aacgctttgg 1140
tagcagcagc aatagcaata gacgacaaat aactggctta aacagaattc aaataaatta 1200
gcatcaggaa gccgcacgag agagtgagtt cgagtttatc agttagttga cttggacacc 1260
gatcatggcg tactgagcgg cgactggcgg cttgttcgtc ggattctgga tgtggagtac 1320
acatagagag agagagagag agagagagag ccagttaata tcgaacctat ccattgatta 1380
taggggggcac ttgcctgcat tatcgacggt ttcggctgct ccgccgactc aatggagaag 1440
cgctgcgttt gactggaatt ggttagtgcc ctcctcaggc ggcgcgattc gatgttgttt 1500
cgcgaattca gcgaacccctt cgataccttc ttgccaaaca gaacggccta aaggagtggg 1560
aaatcggtaa ggtaattagt taaatagaaa aattcagaca ttcgaattat taataaatat 1620
attaataaat atatataaat atttataata cagataaatt tgaaatatag tattgcgcat 1680
actataaccc ccacttccgt tttactcacc ttgaatgcct ctcggaactt gtggctcata 1740
atgttatata acagcgggtt gatgcaggtg gacaagtagt agaggacacc ggagacatag 1800
gtcatcaccg tgtagacaaa ctcgtgctga tcccgcagtt tggccccccg tgcaggggcg 1860
```

tagatggcaa tcagtcgctg ggcgtggaag ggggcccagc agaggaagaa gcacaccacc 1920
acggccacta aaggggaggg gggggcggta aaagtcagtg acgtggggat gaaaatggaa 1980
tggcaatgga aatggggaag acacagcgag agtcacagat aaagagtcaa cgattatatt 2040
ttttggtaat cgcaattgtc atcataaatg gagagtgcca aaatgacaaa cgcagtaaat 2100
tggctgcaga cgggttgatt agcccagctt agggcatatg gtcaaacagt tggcatacac 2160
tgagcaaaag ttgaagtaaa gcttctaaaa tggttcttca ttgcatataa agaaagattt 2220
agatatcata tttcagaatt tttaatttgc aagacttttg tgagaagact aaaccttatt 2280
aaaaacggag tcttctactt ttgtttttgt caaaatttct tactaattta atcatctgaa 2340
tactattaag tcatgtttac ttaatggtta gcggtttaga gatctgtcat ttaatttgtg 2400
ataccccgt actattgctg gtatttttcg ccagtgtata tcgtgctata attacctagc 2460
atcctgagta ctcgccgggt gccatagtga ttgagccgac ctctcaccga gctaagttga 2520
gcccccgagc cgcccatcaa gcccgctgtc cctgctccac ttcctccgcc gttgaaactc 2580
atagcggtac cggatccacc atagcgatat aggatcgtat cactgggcac actcttcagc 2640
tgctgccgtc tggcaaccga ggcaggaccc tccaccaaag tggatcgata caggtgcaca 2700
ccgataagta ggtacaacac taggataata gacatcggag ccaggaagaa tatgaacgtg 2760
gacagctgga atgagtgctt cactatgacc cgcactatgc cgcattgctc cactcccgaa 2820
tagtgctcga ttccaaattg ggcagcctgc ggaatggccg taactatggc cataatccaa 2880
accaggacga tgatgcgaat ggcgcgactg agtttactca tggcctggcc cagaaacgga 2940
tggcaaatgg caatataccg ctccaccgtg aaggccgtaa tcgttagcac cgtggcattc 3000
gccgatgtct ccgccaacag accacgtccg atgcagatgt actccccaaa cacgtacggg 3060
tacttggacc agatgtagga cacctcctgc ggaacgcccg acaacaggag caggaaatcc 3120
gagatggcca gcgaaaagag gtagtaattc gtggccgtgt gcattgaacg gttcttctta 3180
atcacaatgc aggtacttat gttgcccact acaccggtta tgaaaatgag ggagtagacc 3240
accgttacgg gtatcacgat ggccagcgga tcgcgtggcg gtccaagatc atggctcata 3300
ttgccagcgg acatggcgcc ggcggcctgg aatttttccc cttttgtttt cacaaaattt 3360
ccactaacag cgctaaattc cactgccgtt taactgctgc ggctccggct gcgactcgat 3420
gtcccctcaa ctatgctaat tataataatt tattggaggc gaccgggcgg aggagggaca 3480
ggtctggact cgcgcttcgc aatgctagat actcgattct aagattcaga tactcgactg 3540
gcacgcaaac gcaccaaatt accgacgagg tcgagatccg tataccgtat tcagtattca 3600
gtattccgta ttcccggctg tttttctagg gggcccgact tctcttcaat aagttatcga 3660
gttaatttga tttcgattcg aattctgatt cccgagagca gcctgcaacc gcactccagc 3720
gcgctccaaa cttatctgat aagccgcggc gacatttcat tcagatttgt atctgtatct 3780
gaaagccagc gacgaagcgt cgctgttttt ccctcacatg aacgcggaag gtgagcgaca 3840
tggttgtaat gaccgagcgt ccagcgaacg tgccggctca attaaatctt gaaaagcggc 3900
gatatcttga tgggactagc cgatgttaga gactaagcac aatggctccg aatacaacgc 3960
aatttttggg taattaaaaa ggctaattaa caaaaacttg taataatgtt tgctttggcc 4020
catttataaa acacttagtt gggaattttg actttgtgat gcacataatc ctaaataaac 4080
cacgataatg taaacaacac atcttggaat taaaagccaa aataagatgc aataatatta 4140
gatgatgcaa taataccaaa tgaccgaagg ccaccgtgcg tatgagtaat gtccagccag 4200
cttacttgtt agcttataag taagaatatt gttctttgtt tatcttacta aagtatattg 4260
aaatgtaaac aaataataat aattacaatt caaatctaat ttagattaat taaa 4314

<210> 113
<211> 1287
<212> DNA
<213> Drosophila

<400> 113
atgtccgctg gcaatatgag ccatgatctt ggaccgccac gcgatccgct ggccatcgtg 60
atacccgtaa cggtggtcta ctccctcatt ttcataaccg gtgtagtggg caacataagt 120
acctgcattg tgattaagaa gaaccgttca atgcacacgg ccacgaatta ctacctcttt 180
tcgctggcca tctcggattt cctgctcctg ttgtcgggcg ttccgcagga ggtgtcctac 240
atctggtcca agtacccgta cgtgtttggg gagtacatct gcatcggacg tggtctgttg 300
gcggagacat cggcgaatgc cacggtgcta acgattacgg ccttcacggt ggagcggtat 360
attgccattt gccatccgtt tctgggccag gccatgagta aactcagtcg cgccattcgc 420
atcatcgtcc tggtttggat tatggccata gttacggcca ttccgcaggc tgcccaattt 480
ggaatcgagc actattcggg agtggagcaa tgcggcatag tgcgggtcat agtgaagcac 540
tcattccagc tgtccacgtt catattcttc ctggctccga tgtctattat cctagtgttg 600
tacctactta tcggtgtgca cctgtatcga tccactttgg tggagggtcc tgcctcggtt 660
gccagacggc agcagctgaa gagtgtgccc agtgatacga tcctatatcg ctatggtgga 720

tccggtaccg ctatgagttt caacggcgga ggaagtggag cagggacagc gggcttgatg 780
ggcggctcgg gggctcaact tagctcggtg agaggtcggc tcaatcacta tggcacccgg 840
cgagtactca ggatgctagt ggccgtggtg gtgtgcttct tcctctgctg ggcccccttc 900
cacgcccagc gactgattgc catctacgcc cctgcacggg gggccaaact gcgggatcag 960
cacgagtttg tctacacggt gatgacctat gtctccggtg tcctctacta cttgtccacc 1020
tgcatcaacc cgctgttata taacattatg agccacaagt tccgagaggc attcaaggcc 1080
gttctgtttg gcaagaaggt atcgaagggt tcgctgaatt cgcgaaacaa catcgaatcg 1140
cgccgcctga ggagggcact aaccaattcc agtcaaacgc agcgcttctc cattgagtcg 1200
gcggagcagc cgaaaccgtc gataatgcag gcaagtgccc cctataatca atggatagct 1260
gcggacaacg gctggaaaaa ctcctaa 1287

<210> 114
<211> 428
<212> PRT
<213> Drosophila

<400> 114
Met Ser Ala Gly Asn Met Ser His Asp Leu Gly Pro Pro Arg Asp Pro
1 5 10 15
Leu Ala Ile Val Ile Pro Val Thr Val Val Tyr Ser Leu Ile Phe Ile
20 25 30
Thr Gly Val Val Gly Asn Ile Ser Thr Cys Ile Val Ile Lys Lys Asn
35 40 45
Arg Ser Met His Thr Ala Thr Asn Tyr Tyr Leu Phe Ser Leu Ala Ile
50 55 60
Ser Asp Phe Leu Leu Leu Leu Ser Gly Val Pro Gln Glu Val Ser Tyr
65 70 75 80
Ile Trp Ser Lys Tyr Pro Tyr Val Phe Gly Glu Tyr Ile Cys Ile Gly
85 90 95
Arg Gly Leu Leu Ala Glu Thr Ser Ala Asn Ala Thr Val Leu Thr Ile
100 105 110
Thr Ala Phe Thr Val Glu Arg Tyr Ile Ala Ile Cys His Pro Phe Leu
115 120 125
Gly Gln Ala Met Ser Lys Leu Ser Arg Ala Ile Arg Ile Ile Val Leu
130 135 140
Val Trp Ile Met Ala Ile Val Thr Ala Ile Pro Gln Ala Ala Gln Phe
145 150 155 160
Gly Ile Glu His Tyr Ser Gly Val Glu Gln Cys Gly Ile Val Arg Val
165 170 175
Ile Val Lys His Ser Phe Gln Leu Ser Thr Phe Ile Phe Phe Leu Ala
180 185 190
Pro Met Ser Ile Ile Leu Val Leu Tyr Leu Leu Ile Gly Val His Leu
195 200 205
Tyr Arg Ser Thr Leu Val Glu Gly Pro Ala Ser Val Ala Arg Arg Gln
210 215 220
Gln Leu Lys Ser Val Pro Ser Asp Thr Ile Leu Tyr Arg Tyr Gly Gly
225 230 235 240
Ser Gly Thr Ala Met Ser Phe Asn Gly Gly Gly Ser Gly Ala Gly Thr
245 250 255
Ala Gly Leu Met Gly Gly Ser Gly Ala Gln Leu Ser Ser Val Arg Gly
260 265 270
Arg Leu Asn His Tyr Gly Thr Arg Arg Val Leu Arg Met Leu Val Ala
275 280 285
Val Val Val Cys Phe Phe Leu Cys Trp Ala Pro Phe His Ala Gln Arg
290 295 300
Leu Ile Ala Ile Tyr Ala Pro Ala Arg Gly Ala Lys Leu Arg Asp Gln
305 310 315 320
His Glu Phe Val Tyr Thr Val Met Thr Tyr Val Ser Gly Val Leu Tyr
325 330 335

```
Tyr Leu Ser Thr Cys Ile Asn Pro Leu Leu Tyr Asn Ile Met Ser His
            340                 345                 350
Lys Phe Arg Glu Ala Phe Lys Ala Val Leu Phe Gly Lys Lys Val Ser
        355                 360                 365
Lys Gly Ser Leu Asn Ser Arg Asn Asn Ile Glu Ser Arg Arg Leu Arg
    370                 375                 380
Arg Ala Leu Thr Asn Ser Ser Gln Thr Gln Arg Phe Ser Ile Glu Ser
385                 390                 395                 400
Ala Glu Gln Pro Lys Pro Ser Ile Met Gln Ala Ser Ala Pro Tyr Asn
            405                 410                 415
Gln Trp Ile Ala Ala Asp Asn Gly Trp Lys Asn Ser
            420                 425


<210> 115
<211> 5927
<212> DNA
<213> Drosophila

<400> 115
ttaagccaca aaatgattga aaggaaatgt tgtggggtaa aagtaaaatt ctcagatcga 60
tgagtacaat gtactatgta cttccctcgg cagctcctgc tgtgttcctt cagccgtttg 120
tgtctgtttg tccctaaagt aattaaaagt aagtaaagca aagtcacccc ctcgtcggtg 180
ggcgtggctg tccagtgatt tgccacccgc aacgattttt ttttccagta ctttcctttt 240
tttctttttt ttgcgtttat gcaaatggcc gctgaccaaa agaattaatg cgtcccgaaa 300
gtgagttgaa ttggaatgag ccgcaattct cggagtccaa aaactgcggg cagcctgtca 360
gccgtctgtt ttcacttctt ttgcgggcgg gacgaaacat caacgaatgt caatgaattt 420
acaggggaat gctataaaaa gcccgaccaa acaaaaaatg taaagcccgt aattccccat 480
tgtctcagtg cgttgctaat ttggccatat ctttaaatag ccagccgaac ggaaaggaat 540
catttgtaat ttaaagtcca attattgtgc catttgaggg gaaaactgtg gcatgccagg 600
gaatggaatg cctcgtaagc cacatcgctg cacttcatta aattgacgag acaaaagatc 660
tctacacata caataatgga aaatatttgt tccacaataa aaggggttga ggaggcactt 720
tgtgttaatg cccaacagct gcaaatctcg gtctgcggcg gccttaaaaa tagagaaatg 780
aagatggacg gtccgatgat gatccgcttg gctgaaatta ttctgtaaat tctacctatt 840
tatatcgtaa atatagattt tgatgctgat gccgatgctg cgcgagtaaa accaaattcg 900
gcccacgcgg cgtatgcgcg tttttcgcac ttaactcatt aaaacgtttg ccgctggccc 960
attgaaattt aatgggaaaa cacgcgctgt tttttttttt ttagaaattc gcctcaatag 1020
attaagctcg taaattattt aattgcattc ttgatttttg atttgaaatc aaaacattcc 1080
ctcatttaaa aactcattag gaatagcaat ggaaacgtgg ttgagcggat gtttatattg 1140
tagaaaaaaa aaatatatat atatcaaata tatgaaaaag gttttgaagt ttgtcagtca 1200
ataaatcaat ataaagcccc acttaaaatt acgacacaaa agtttccatc aaacgaaatg 1260
caaacgaaaa caatgctctg aaaatatttc atcgatatat tgaccgagca ctcagatgtc 1320
gtcacttgaa ggacccttct cgtccagccc ccgagtgcac caacccccact caccttgtag 1380
aaggccttcc ggaaattctc ggagaggaag gcgtagagca gcggattgat acacgagctg 1440
ctgtaggcca gagtctgggc ggtgacctgg atgactagct tggtgagggt gttcgtctcg 1500
atgacatcca gtgacttgag cagcaggatg agctgttcaa tcaaacgaaa gaatatgttt 1560
gattatttaa aatatatgtg ggtttcatat gtattattgc gtctcacctg gacaggcagc 1620
cagagcgagg cgaaggcgat gaccaccacg acgacgagtc gggtgacccg cttgcgaccg 1680
cgctgcgact ccttggacat gcggacgccg gttccctggc gccagagccg catgatcatg 1740
cgcatgtaga gaccgctgat gatcatcagg ggcagcaggt aggagctgat gaagaaggtg 1800
acctggtagg tgcgcggacc aaggaagtcg ttcgtcgtga aggtgcacat gccgtaggtg 1860
atgttcttct ttgcatcgta atccacctgc ggatatcgat gagaatgtgg ttgtgggaac 1920
tatgaagaaa cgtgtgcggt gtgggaatcg aatattgcgc atacgtcatg ttggcggtga 1980
cagttgcatg gcagccaagc aacgtcaaac ggccaacagc cagacatcgt aaataaaact 2040
aaactgtcat acgaatagtt cggacgtcga caactaactc gcaatctcac tttgtttctc 2100
tgtcgctgac atttggctga caaattgttc agccggagtt ttggatttgg ataaaagcgc 2160
aatgggtatg ggcaatcggg ctttcgaggc ttgttgtttg ggggatggag aatctattgg 2220
tctgttggct agacgagact ggtttgctga cgcttaggag ggtcgtaaac tgcccggtga 2280
atcgtaaaaa tattcgcagc cttttgttta catttttgtt tacaatctcc gccccttcac 2340
cgcgctttca atcgatgaat gcgtgtggaa aattgcattt aaaaaatctt cccattaatg 2400
```

ccggcatgaa ttgcttacaa aaatgtttgt ttatggcccg cccttgtttt cgttacactg 2460
agcgagcgat gttggcccac aaattgagtt cgaactctga atctgcgcag cgtcaagtgg 2520
ccaaggccac agaagtagac aaatcatggc cagaccctcc gctgatttat gggttcattt 2580
cgctggcatg ttctccgcaa aacccaacga cttcatagac ttgacaagcc attaatgccc 2640
tttcctcgtt cgcgagaaat cagccaacag ctgtcaatca atgccatgac ctggagatcg 2700
catatctagt tacgcagtcg gacatatgtg catttatgat tattatttga tatcgatttc 2760
taataaaaag ataatggcac acgctctcac aggctcattt gatatgcata taccaagcat 2820
gtagctgggc aatgtaatat ggaaataatt aatagcaaac tgttaattaa catgttgatc 2880
aaacatgact taattatgcc tttttatcgc ctgttccttt gtatttgcac acgcagacac 2940
cagataacaa agcaaagcac acatatgagc attcaatgtg ggctctattt ttccagctta 3000
atccttagac aaagcaaagc aaacacataa acaaacaata atgccaccca ttggagcact 3060
ctccccaatg cgaacaaaaa ccgaacgaga caaaaataac aatataacga agcgtgaaat 3120
gccatttaat aaaatattag gttacaataa caacaccggg agcaacagtg ggccattgcc 3180
atctcacaga taaaaaaagt tcatcatccg tgtgaagatg agattcgttt aattaaccgc 3240
gactgtcgct tttggccaag gaaaagggga aacacaacgg cacaaacaaa aagttggtca 3300
tgtgtgatgt gggaatcgct ttattggcac tgggaaatca gtggaaaaag cccgcgggcc 3360
atttatgagt atcaacagaa attggaacca atgtccgacc ataattacgt gtagttgata 3420
ctcgcaatga aaaaatatga tagatattca gtattaaata gcactcaaca attgtttgaa 3480
ctgtaccgga aaattgaata ttagttaata atcccagtgt tttcttgcga ataaaaactt 3540
aaggttaaag tctgcaaata caaatacatt gtatatcctc aaaatcgccg gtaaattccc 3600
cttttcattt ttgatatcac ccgctgatct cgagcaaaac aattccgatc ttcgtcgatt 3660
attccgcaat taaattgttt ctcgcacttt tcgcatcgat cataaaagcc taatgtccaa 3720
gatgagcatt tcaaggtcgc cacccgcgaa gacatttatc atacgcagtg tgcgcccaaa 3780
gttcgcaagt gcaaagtttc cataatggcc gcttgttacg acttacgact gcccatgatt 3840
tacgcgcttt tggatcaaca gttgaatttg aaattttgca agccacacgg agaggaggtg 3900
ggcagatatg ggcagatacc gcccagttgg caagtgtttc tcgctggcta tctggtctat 3960
aattaaaaat ttatgaattc aaaatagcaa attagtaaca aaaacaaatc gggcgccgcg 4020
catcgcgaaa ttttcaatat aaacacaaaa cagaaaataa gataaaatgg ggatacttta 4080
ataggggcat atagtaaaaa caaatgagtt gtatcaggat cagtcgattg gacgatcgct 4140
ggtttgtgcg gggtcttcgc tttctggatt acgatctggc agaaggcaga ccatctatct 4200
aggtatctat agctacatta gaaccaggca gtcggcacca ttgacgtgca gcgagaggaa 4260
tctaacgagg tgacgatgat ggaaagcaat aaatttgcct ggcccgtgtt cttcacagca 4320
aattcaggca cagcaggccc acttatagca ttctacccac tcctcccact cttccactcc 4380
cacttaccac cacgtcgtgg gtgaaggcca ctggcaccga aacgaccagc accacgatcc 4440
acagagtcac gatggcaatc agggtaatgt tctccgtcct catcatccgc gagcgaatgg 4500
gatgaaccac cgccaggaac cgatcgatgg acatcagcac cagcgtgtaa atggaggcga 4560
aggcggtcac cacaatcagg tactggacac tgcggcacca gaaccttcca tatggccagt 4620
agtacaccat gtaatcggtg gccgtgaagg gaatgcagag gattacgaac atcagatcag 4680
cggcagccag attgacaatc atcaggttgg tggtggagcg catgttgttg ttgaagacca 4740
ccaccaggat gaccagcagg ttgccgaaga atcccgtgat ggcgatgacg ccgaagaaga 4800
atcccactat ccacggcaaa gtgccattga tggccagcca ctcttcgtcg gtgaagaatg 4860
aggtgatatt ctcctcattt ctggttgcat ttaggctaat attagccagc attgtggtgt 4920
tctccatgtt ttctcagctt tagtgcacaa ctttccggtt cactttgtac tggctattgg 4980
cgatttcgct tgattgtcac gggacatttt caattttggg ctctttttcg ctcacagttc 5040
atcgaacgca aataacaaat ttccaattag gagagtttca caaagcagcc gccacaaaaa 5100
ttcaattatc atgcggaata ctacgaccgt tgagcttcgg agacctcaaa aacgtttcgc 5160
atacgccgtg tggtatcgaa ctcgaccgga attgttcgca acgtggcggg ggcaaaagca 5220
atttggtcct tccacgataa cggtaattgt aggtgggtag cacaggtaag acaacaataa 5280
aggacagcac acaaggatca cggtgaggcg aaaggcaaga gagtgagttt aaaggtgagg 5340
gcacgtgcaa acgacaccga cacagccaaa ctgtgccgag tcacaccctg acgacaagga 5400
caccggggtg tccttaggtt cgaggaccac gttgctgcac ctacaccata aattcaatat 5460
gatgtcttcc tttcgctgcg tttctgttta catttctcgt ttgctttttt attgacaaac 5520
gaacggcata aattttgggg ctataaggct atggccaaaa aaccacgttc ctattgttgc 5580
taccgactgc tgcgccgcgg tacgctatat attttcgatt ttatactgct gcttttggcc 5640
aattcgatag tccacactgc tcgcgtccgt atccagcgcc caactgattg ctggctgagc 5700
ttgttgagga atacttcgat gcctcggcgg cggcaaaacg ccggcagcga cgccagccat 5760
cccagcaaga ataatggcaa caatagcaac aaccagcgca gctggcatct gtgcgtgtga 5820
tggttttggg tggccgccgg agccacctct tcctcctcat cctcctcatc cgatcccaat 5880

ccccatcagc tgggctggaa tgtttgctct tgttgctccc accgcta 5927

<210> 116
<211> 990
<212> DNA
<213> Drosophila

<400> 116
atggagaaca ccacaatgct ggctaatatt agcctaaatg caaccagaaa tgaggagaat 60
atcacctcat tcttcaccga cgaagagtgg ctggccatca atggcacttt gccgtggata 120
gtgggattct tcttcggcgt catcgccatc acgggattct tcggcaacct gctggtcatc 180
ctggtggtgg tcttcaacaa caacatgcgc tccaccacca acctgatgat tgtcaatctg 240
gctgccgctg atctgatgtt cgtaatcctc tgcattccct tcacggccac cgattacatg 300
gtgtactact ggccatatgg aaggttctgg tgccgcagtg tccagtacct gattgtggtg 360
accgccttcg cctccattta cacgctggtg ctgatgtcca tcgatcggtt cctggcggtg 420
gttcatccca ttcgctcgcg gatgatgagg acggagaaca ttaccctgat tgccatcgtg 480
actctgtgga tcgtggtgct ggtcgtttcg gtgccagtgg ccttcaccca cgacgtggtg 540
gtggattacg atgcaaagaa gaacatcacc tacggcatgt gcaccttcac gacgaacgac 600
ttccttggtc cgcgcaccta ccaggtcacc ttcttcatca gctcctacct gctgcccctg 660
atgatcatca gcggtctcta catgcgcatg atcatgcggc tctggcgcca gggaaccggc 720
gtccgcatgt ccaaggagtc gcagcgcggt cgcaagcggg tcacccgact cgtcgtcgtg 780
gtggtcatcg ccttcgcctc gctctggctg cctgtccagc tcatcctgct gctcaagtca 840
ctggatgtca tcgagacgaa caccctcacc aagctagtca tccaggtcac cgcccagact 900
ctggcctaca gcagctcgtg tatcaatccg ctgctctacg ccttcctctc cgagaatttc 960
cggaaggcct tctacaaggc gaatttctaa 990

<210> 117
<211> 329
<212> PRT
<213> Drosophila

<400> 117
Met Glu Asn Thr Thr Met Leu Ala Asn Ile Ser Leu Asn Ala Thr Arg
1 5 10 15
Asn Glu Glu Asn Ile Thr Ser Phe Phe Thr Asp Glu Glu Trp Leu Ala
20 25 30
Ile Asn Gly Thr Leu Pro Trp Ile Val Gly Phe Phe Phe Gly Val Ile
35 40 45
Ala Ile Thr Gly Phe Phe Gly Asn Leu Leu Val Ile Leu Val Val Val
50 55 60
Phe Asn Asn Asn Met Arg Ser Thr Thr Asn Leu Met Ile Val Asn Leu
65 70 75 80
Ala Ala Ala Asp Leu Met Phe Val Ile Leu Cys Ile Pro Phe Thr Ala
85 90 95
Thr Asp Tyr Met Val Tyr Tyr Trp Pro Tyr Gly Arg Phe Trp Cys Arg
100 105 110
Ser Val Gln Tyr Leu Ile Val Val Thr Ala Phe Ala Ser Ile Tyr Thr
115 120 125
Leu Val Leu Met Ser Ile Asp Arg Phe Leu Ala Val Val His Pro Ile
130 135 140
Arg Ser Arg Met Met Arg Thr Glu Asn Ile Thr Leu Ile Ala Ile Val
145 150 155 160
Thr Leu Trp Ile Val Val Leu Val Val Ser Val Pro Val Ala Phe Thr
165 170 175
His Asp Val Val Val Asp Tyr Asp Ala Lys Lys Asn Ile Thr Tyr Gly
180 185 190
Met Cys Thr Phe Thr Thr Asn Asp Phe Leu Gly Pro Arg Thr Tyr Gln
195 200 205

```
Val Thr Phe Phe Ile Ser Ser Tyr Leu Leu Pro Leu Met Ile Ile Ser
    210                 215                 220
Gly Leu Tyr Met Arg Met Ile Met Arg Leu Trp Arg Gln Gly Thr Gly
225                 230                 235                 240
Val Arg Met Ser Lys Glu Ser Gln Arg Gly Arg Lys Arg Val Thr Arg
                245                 250                 255
Leu Val Val Val Val Val Ile Ala Phe Ala Ser Leu Trp Leu Pro Val
            260                 265                 270
Gln Leu Ile Leu Leu Leu Lys Ser Leu Asp Val Ile Glu Thr Asn Thr
        275                 280                 285
Leu Thr Lys Leu Val Ile Gln Val Thr Ala Gln Thr Leu Ala Tyr Ser
    290                 295                 300
Ser Ser Cys Ile Asn Pro Leu Leu Tyr Ala Phe Leu Ser Glu Asn Phe
305                 310                 315                 320
Arg Lys Ala Phe Tyr Lys Ala Asn Phe
                325
```

<210> 118
<211> 7383
<212> DNA
<213> Drosophila

<400> 118

```
atcttttaag taaaaagcgc tcaagtcaaa acttttactt tcatttaaaa gtaaattttc 60
ttgatgtttc aaatttatag caaaatgcca taatttcagt tcgtttgaaa caaacgaaat 120
gaagcggaaa ttgcgaaccg aatttattcg tacaatttat gggccacaca tgaacaataa 180
taattcattc ggctgcatgt aattaagaat aataaaaata ctggtatatt ttgtttgttt 240
gcttaatgac tcgaaatgtt tattcgtttt cgtttggtaa ttaaatacaa taataattac 300
attaataagc ctttctgttt tttaatgaca aatattgtac attacatttg cttacatcca 360
gctgtgtgct tgtgtaggat ttaatacata atcggattgt tgttaaaata ctatttaaat 420
aatttaattt aatgctactt aattgatttc aggtgctaat ttaatttgat tcattttttg 480
ttcagctaat aatacatttg atttggtgag ttttggtttg agtttgataa ggtcattaac 540
gaatcgatta tttacaacgt ttggacttcg tgctttaagg atcgcaattt gaagatacca 600
ttattataat actgtacacg gctcacgacc agcattcaca ctcattcaat atatatcctt 660
ctgcgaaaca attatcagca atgggacttg taatcacaat tcattaagac gaaatcacag 720
aggcaaatgg tcatgttcca atgacacatt tagttgcggg atatctaaac actgtaaagg 780
ggttttacta aagccttaca tacaattaca ataaaattta aacatattag tgtagtagct 840
ataccctaat aatagtcatt cactaaaaca ttaacacaca tacgttaaat gaatgagcta 900
tgacagccaa tttactggaa tacacatcat atgtgaatgc cgactggctg cggaatgtgt 960
cattggaaca cggctataac ttagatcgct ctagtttaaa ttaaagtggt tgccacaagg 1020
acttgcccgc atccttgtcc ttcgcctgcg agccgggcag gtggactcgc gacacctcgc 1080
tgagtatgca ggtgccatcg cattcggttt ggctgccaaa ttcctgttcg caaatgcagc 1140
tggagctctc cctgcttagc agggccaact ggactccggc cccggatgcc aactcctcgc 1200
tggggggcgt ggttgatggc agctccatga gtgccacgtc gtcctcgcat tcgattacgt 1260
tcggatgcag aggtttgaag ttgtttcgcc gccacggcgg cattgcaacg ccaacacctg 1320
gtgccgaaac acccggagtg gccgcaaagg taaccacgct ttgatgctga tgatgttggc 1380
cattgacatt gttgttctgc ccactacttc cgttgtaaat gccgctgtta gcaccggatc 1440
catgcacccg cggcatatgc aagcccggct tcccattgtt gacaccacca cgcgccggac 1500
caaagagatt actccggatt cgcatcgagg agttggcgta ggttgacctt atggagctgg 1560
cgcgggtgtg catcgaaaag gtcctttcgt gggcgtccat gctcgtcttg aacttgcaga 1620
aacaggccgc aaatcgcttg ttgaattccc gcttaaattt ttcctgtgca gagaaaaata 1680
gacgacggtt taggttaatc catcccaata acatatcact aaagtaagcg aaagtacttt 1740
tattcactta aatacaaatt aaagtgcatt gcataaatca taaacctctt agcaaagttt 1800
atcatttgca aaaatacatt ttttcgctat caaactttct ggttaattat tatgtgaatt 1860
ggcagaaaat ttttcgtgaa aaagcgaaaa ctgccccggc tacggcatct ttctggctaa 1920
tgaaagtgct ctaggcaaac ctccatcttc gctgacaaat cccatttgtc ttcgctgaca 1980
gtaaaataga atatctgttg acaaagtttt tcacgcacag ccagcagggg aaaaagttta 2040
cgagtttatt cggagatctg gctctcaatc tatcttttac agacaatttg agtgtcaaag 2100
caaacagccg acgggttttt tcccgctttt atcggtttgc cactttggca ttctgaggat 2160
```

ttttgtggat ttcgcttcag ttttgcggtt tcacaatttc tgcggcccag acggtaggtg 2220
tttgcttatt ggctacagtg gagcgctgct ttaattaact ttgacaaagt atttttttgt 2280
gtcagctatt ttaggtgcga ggcgaaaaga aattggaaat tgtttcaaag taaggcagaa 2340
gttattttg attaaaaaaa aagtgacgcg tatgcttata aatcagcgag aaggaaaatg 2400
ggaaaacatc aacaaagcaa tcgattcgct tacaatgaaa ttaagttaga catatttcat 2460
tctttttttaa gcttctttga tttttgagac catgaaaact tccgtaaaag tttcatttgc 2520
tttgcgtcta tgtttgtgtt tcatttgatg agataataga aaaagctaat tttggaagca 2580
cgtgcaaagg ctttttttct tttaggtcta aaagtacgaa aataatatag aaaaaattac 2640
tttttttact ttaacttcat tattaaaagc aagacgttaa gtatgtagat attttgacga 2700
gacatccaat tgtacgcaag gcaagttttc ccaagatccc ctgtcaataa ctgactcacc 2760
caaacttaag ttccccaaaa ctaacccaag gcgaccttga gtccctcatc aatccccaca 2820
tttccccggc aaatttacca cacatttata gacttatgtc tttttagtgc gtgtgcaaag 2880
aacgtgtccc tgagcacgaa caaacatcca aaggatgtca atcaaatccc agccagacaa 2940
atgcaaattt aacactccgt aaacaaatca tcatcatcat tggagcccac acttacattg 3000
tagatgccat aaataaaggg attgtagcag ctattgctca tggccagcca atcgcagcaa 3060
aaccagacga tgctaatgaa gtggtagtcg ttgatttccg gtatcgtgac atacagaata 3120
ttatagagct gcagtggcag ccagcagagt ccaaagataa tgaccacgat aatcagcatt 3180
ttgatgacct gcaagcgtcg cagaaaaagg aaattaaacc cagtcccgat ttgcatacca 3240
catcaacgtg caaaactcac cttcttcttg tttttcaaca gcgttatgtc ccgtgaatcc 3300
tgtgcgttac caggagcacg tgtgccccac aatcgtaccg ccatctggat gtagacaaag 3360
ctgatgacac agaatggaac cagatactgc acaaaaacca gggtgtagcg aaaggattgc 3420
aattgatcat cggataggtt cttgttcatg cagaatggcc gcgtcacatt gtaggtctca 3480
ttgttctctg ttcaaaggat gaaattcagt taaggtcgct ggcacttaat ccatgcaagt 3540
tgacgtattg acttaccgcg aaatctttcg gtcaactcct ccacacgaaa ggcaatggca 3600
aagggcaccg caaatagcag ggccagcatc caaattccac caattatgaa cttcgatacg 3660
aacttggtgg gacgtgccct gcatgcgtaa gccgattttt aatgattatg attgacgaat 3720
cgaatcgttt cgaatgtggg agcccactct cccccccagc gactcaccta agtggattaa 3780
tgatggccct atgccgatcg attgcaatgg cggtcagcgt gaataccgag acatttacac 3840
tcagggcctg gacgaagggg cagaagctgc acatgaacca cggcaggttc caactctgca 3900
gcagggcagc ctggaactac acggaatgaa aatattgcaa ttattagttt aaagtattgc 3960
agattatcag aaacaacttg actatacaac tatgcctaac ctttagcctg agcatgatgt 4020
ttacaaatca caaataatcc ctttaaaatg ttcactatat atatattcat tttaattaaa 4080
gtgtaaggaa ataggatagc tacaagtgag tcgtccgcct gaggtttgat atgatttaat 4140
caacccaaaa ttgcgactcc attcaagtca cacggcaaac accattaatg aacagtggaa 4200
catgtcgggg gtgttgtatt ttcggccagt ttcaacattt cgtcgacaaa taaacgctcg 4260
aaagcccagc ggtgttgttt tggacactga tttgcgcagg agtgagtgaa aaagtgtgtg 4320
tgtgattact gcggatattt gaagaccgcg ctttctccgt tctccacaca aagtcggccg 4380
caaattgcaa tttctcgccg gcttattgat gacacttact ggcggctgtg acaggaccaa 4440
aaatctattt gattgcggca ttaatataaa ttaagtgccc aaaatgtggc tggctttggg 4500
tcaagttcaa ggttcaacga agtgtacaga ttgagcttta ttagtatgaa ttctgaggta 4560
aaaatcaatt tactattttt atttgaaaca tagaaagatg taaacaattt gtactaccaa 4620
ttaataatta tttcgactat tcagttcgat ggttacattc aatattttac gtctaaaaca 4680
agctaaaatg caattgcggc attctgcgaa acgagctgaa tacaaagtaa cgattcaatt 4740
actttattca aatgtccatt gctcattgat tgaccatttt ccgactaaac agtattttca 4800
aagcaaactg cgttcgcgtt ttgcattcga cgacaccagg gaatcaattg aagataatgg 4860
caattcaatt actcggtatc acatacaaat tccaaagaaa gcccagaaaa ctgttcacat 4920
agcttttagc actctatttt gggtacttta tagaatacaa atcccattaa cacataaacc 4980
gaaaatcaca tttaaatttg atgtaggtaa tattccatta ccacaacttt tttcgacaag 5040
atgcgcgtag gccatggcaa ttttctaatg ccaaacatga gcgctcgttg cccggaaatt 5100
aaattgatgc cggattatgc tgaaattggt gtccctgcgc actgtgggtt tcaaaaactg 5160
ggctggctac ctttctatgg tgaaaataag cgaaaactaa tgctaatgtt gttggcttgt 5220
gtgatttatg cgttaaagtc gatatccatt tactagcgca acagtgcgta tgggcaattt 5280
tgcatgtatg catttcatat ccgcgaaaaa aaaaatctcg gtaattaccc gcattataat 5340
taagcttaac tgttgcaact cacctgaaat ggtatgcaga agaggccaat aatcacatcg 5400
gcaaaagcca aattagcgat atacatattt gtcacggtcc gcatttgcct ggtcgtggcc 5460
accacccaga tgaccaaagt gttgccaatg acggccacga tactgattcc cccatagaaa 5520
atggacaaca gggccacaat ctcagcggga gccgcgtata gacgctctgt gaatgaattg 5580
aaaaatagtt gtgcattaga tttatatcac tcggaaaaag tgcttgcaac attattataa 5640
aactaaatga gatataacaa tctgttttac acaaaaattt aagaatgtat ccaatacata 5700
atgcgctcct actaagttac tgaaatttta gttatttatt acaagaccat gcttttatga 5760

```
gtgaataaat tttgtgtgtg gtgcttagct tcttccaatc cattagtttc agtgttgaac 5820
ttgtgaaaaa tcgcgcttgc tgacagaagc agattaattc aagtgtcagt tgtcagtcaa 5880
ctgatggcag gacagacgac attcacgttc ctgctgactt gtaaacgtca aggaaattga 5940
accgtatttt atagccaaag agggaatttt ataaatgcca ctttggacag tatatacctg 6000
tatatatata tacctatatc caatttaata taaaatattg atagttaaac ccaactttga 6060
tatgaatttc attagcatat ccgaacatca cctgtttcct catatcaaag tcaatgagat 6120
ggttgagccc aagtttggtc tttaaattcg gcacccccca cactcatgaa tgattcaaga 6180
tcaaagaact aacggcgaac tgtcagaaga aaaaaacact tgtcaatgtt tttgattgat 6240
gatacgtcaa gtcaatttaa tccagttgga tgagtcgcgt ccccagttgc ttgaagggat 6300
tagagagtat acttaccaaa ttctgcttcc tcctcggctc cgggcaggaa ttccaggcgg 6360
gactcctgct cgattaagtc cattgccata acactttttc gtttttagt tttagttgtg 6420
tatcctaatg aggatatgat atccttcaaa ccttgtattt tagttcgaat gctgctgctc 6480
tatgtagtgc aatacatctt caggggcttc tgagcttcat tgcaatctaa actggaaaga 6540
tataaattac agcatttgtt catttgtttt ctgaaaattc gcaaaaaagt tatctgccac 6600
aagtccagtg gtcagggaac aacaaaagtg cagacaacaa gagaaacaaa tgatctgaaa 6660
aaaaatcatt gaaaagctat acaagcaata tcagtcttgc cctcaaggaa ctgaccaaat 6720
tcagtgggcg gacatgagtt atggctcagg tggaatcacc tgtcagcttt cacaagtgaa 6780
aagtttgttc tgtataattg aaagtgcacc caaattcagt ttggcgactg agggcactgc 6840
actagaattt cgcttgtatt tcgtttctta tgaaaacagt tcatgttttc atgtttttat 6900
tgaagagtca tgccaacgct caaatgttgc cttagcctta ttgaggtggc tcctaaaaaa 6960
taaacaagaa agtcagcaag ttcattgaat gtttgagaat agaagcctca ctcaggttaa 7020
ttggccattt atcgaactgt taacattgtt gggacttgtt aatttatctt cttcagtgta 7080
aacactgtca agacttgtta atttattggg taactaaatg tttattctat aggcagacgg 7140
aattgttata gattgactta gggatgcttg taggaaatag gctgggaaat tggcgaatgt 7200
cgtgaaaaaa agatgtattt agttttgaat gactgtacta tttcccccag aaagaaacag 7260
gttcacaggc tttaagttta tatggaaatg caaaaatttg tttcgagcat taaattactc 7320
attttattat gctggtcaca agctaaaaga actaaaagga gccaccctaa cgcatgaccc 7380
caa                                                               7383
```

```
<210> 119
<211> 1623
<212> DNA
<213> Drosophila

<400> 119
atggacttaa tcgagcagga gtcccgcctg gaattcctgc ccggagccga ggaggaagca 60
gaatttgagc gtctatacgc ggctcccgct gagattgtgg ccctgttgtc cattttctat 120
gggggaatca gtatcgtggc cgtcattggc aacactttgg tcatctgggt ggtggccacg 180
accaggcaaa tgcggaccgt gacaaatatg tatatcgcta atttggcttt tgccgatgtg 240
attattggcc tcttctgcat accatttcag ttccaggctg ccctgctgca gagttggaac 300
ctgccgtggt tcatgtgcag cttctgcccc ttcgtccagg ccctgagtgt aaatgtctcg 360
gtattcacgc tgaccgccat tgcaatcgat cggcataggg ccatcattaa tccacttagg 420
gcacgtccca ccaagttcgt atcgaagttc ataattggtg gaatttggat gctggccctg 480
ctatttgcgg tgccctttgc cattgccttt cgtgtggagg agttgaccga aagatttcgc 540
gagaacaatg agacctacaa tgtgacgcgg ccattctgca tgaacaagaa cctatccgat 600
gatcaattgc aatcctttcg ctacaccctg gtttttgtgc agtatctggt tccattctgt 660
gtcatcagct ttgtctacat ccagatggcg gtacgattgt ggggcacacg tgctcctggt 720
aacgcacagg attcacggga cataacgctg ttgaaaaaca agaagaaggt catcaaaatg 780
ctgattatcg tggtcattat ctttggactc tgctggctgc cactgcagct ctataatatt 840
ctgtatgtca cgataccgga aatcaacgac taccacttca ttagcatcgt ctggttttgc 900
tgcgattggc tggccatgag caatagctgc tacaatccct ttatttatgg catctacaat 960
gaaaaattta agcgggaatt caacaagcga tttgcggcct gtttctgcaa gttcaagacg 1020
agcatggacg cccacgaaag gaccttttcg atgcacaccc gcgccagctc cataaggtca 1080
acctacgcca actcctcgat gcgaatccgg agtaatctct ttggtccggc gcgtggtggt 1140
gtcaacaatg ggaagccggg cttgcatatg ccgcgggtgc atggatccgg tgctaacagc 1200
ggcatttaca acggaagtag tgggcagaac aacaatgtca atggccaaca tcatcagcat 1260
caaagcgtgg ttacctttgc ggccactccg ggtgtttcgg caccaggtgt tggcgttgca 1320
atgccgccgt ggcggcgaaa caacttcaaa cctctgcatc cgaacgtaat cgaatgcgag 1380
```

gacgacgtgg cactcatgga gctgccatca accacgcccc ccagcgagga gttggcatcc 1440
ggggccggag tccagttggc cctgctaagc agggagagct ccagctgcat ttgcgaacag 1500
gaatttggca gccaaaccga atgcgatggc acctgcatac tcagcgaggt gtcgcgagtc 1560
cacctgcccg gctcgcaggc gaaggacaag gatgcgggca agtccttgtg gcaaccactt 1620
taa 1623

<210> 120
<211> 540
<212> PRT
<213> Drosophila

<400> 120
Met Asp Leu Ile Glu Gln Glu Ser Arg Leu Glu Phe Leu Pro Gly Ala
1 5 10 15
Glu Glu Glu Ala Glu Phe Glu Arg Leu Tyr Ala Ala Pro Ala Glu Ile
20 25 30
Val Ala Leu Leu Ser Ile Phe Tyr Gly Gly Ile Ser Ile Val Ala Val
35 40 45
Ile Gly Asn Thr Leu Val Ile Trp Val Val Ala Thr Thr Arg Gln Met
50 55 60
Arg Thr Val Thr Asn Met Tyr Ile Ala Asn Leu Ala Phe Ala Asp Val
65 70 75 80
Ile Ile Gly Leu Phe Cys Ile Pro Phe Gln Phe Gln Ala Ala Leu Leu
85 90 95
Gln Ser Trp Asn Leu Pro Trp Phe Met Cys Ser Phe Cys Pro Phe Val
100 105 110
Gln Ala Leu Ser Val Asn Val Ser Val Phe Thr Leu Thr Ala Ile Ala
115 120 125
Ile Asp Arg His Arg Ala Ile Ile Asn Pro Leu Arg Ala Arg Pro Thr
130 135 140
Lys Phe Val Ser Lys Phe Ile Ile Gly Gly Ile Trp Met Leu Ala Leu
145 150 155 160
Leu Phe Ala Val Pro Phe Ala Ile Ala Phe Arg Val Glu Glu Leu Thr
165 170 175
Glu Arg Phe Arg Glu Asn Asn Glu Thr Tyr Asn Val Thr Arg Pro Phe
180 185 190
Cys Met Asn Lys Asn Leu Ser Asp Asp Gln Leu Gln Ser Phe Arg Tyr
195 200 205
Thr Leu Val Phe Val Gln Tyr Leu Val Pro Phe Cys Val Ile Ser Phe
210 215 220
Val Tyr Ile Gln Met Ala Val Arg Leu Trp Gly Thr Arg Ala Pro Gly
225 230 235 240
Asn Ala Gln Asp Ser Arg Asp Ile Thr Leu Leu Lys Asn Lys Lys Lys
245 250 255
Val Ile Lys Met Leu Ile Ile Val Val Ile Ile Phe Gly Leu Cys Trp
260 265 270
Leu Pro Leu Gln Leu Tyr Asn Ile Leu Tyr Val Thr Ile Pro Glu Ile
275 280 285
Asn Asp Tyr His Phe Ile Ser Ile Val Trp Phe Cys Cys Asp Trp Leu
290 295 300
Ala Met Ser Asn Ser Cys Tyr Asn Pro Phe Ile Tyr Gly Ile Tyr Asn
305 310 315 320
Glu Lys Phe Lys Arg Glu Phe Asn Lys Arg Phe Ala Ala Cys Phe Cys
325 330 335
Lys Phe Lys Thr Ser Met Asp Ala His Glu Arg Thr Phe Ser Met His
340 345 350
Thr Arg Ala Ser Ser Ile Arg Ser Thr Tyr Ala Asn Ser Ser Met Arg
355 360 365
Ile Arg Ser Asn Leu Phe Gly Pro Ala Arg Gly Gly Val Asn Asn Gly
370 375 380

```
Lys Pro Gly Leu His Met Pro Arg Val His Gly Ser Gly Ala Asn Ser
385                 390                 395                 400
Gly Ile Tyr Asn Gly Ser Ser Gly Gln Asn Asn Asn Val Asn Gly Gln
                405                 410                 415
His His Gln His Gln Ser Val Val Thr Phe Ala Ala Thr Pro Gly Val
            420                 425                 430
Ser Ala Pro Gly Val Gly Val Ala Met Pro Pro Trp Arg Arg Asn Asn
        435                 440                 445
Phe Lys Pro Leu His Pro Asn Val Ile Glu Cys Glu Asp Asp Val Ala
    450                 455                 460
Leu Met Glu Leu Pro Ser Thr Thr Pro Pro Ser Glu Glu Leu Ala Ser
465                 470                 475                 480
Gly Ala Gly Val Gln Leu Ala Leu Leu Ser Arg Glu Ser Ser Ser Cys
                485                 490                 495
Ile Cys Glu Gln Glu Phe Gly Ser Gln Thr Glu Cys Asp Gly Thr Cys
            500                 505                 510
Ile Leu Ser Glu Val Ser Arg Val His Leu Pro Gly Ser Gln Ala Lys
        515                 520                 525
Asp Lys Asp Ala Gly Lys Ser Leu Trp Gln Pro Leu
    530                 535                 540


<210> 121
<211> 4391
<212> DNA
<213> Drosophila

<400> 121
agtggaaatt tcgaaattga aatggaccct ttccgccgaa aaatcgattg attttcaatt 60
tcccctctgc gctttcttgt tttgctaagt ttgaaagtta tttgctttaa ctcatgctga 120
aagaaaatat cagttgacta gccgcttttt aacctttttat tcgaatttta tccttagttg 180
ttgcacaaaa gagtctatgt gcttagttga tgatatcagt cgatatctct ttaattggcc 240
aataactgct gaatatgtgg cagcacaata tagttacgaa agtaatcggc ttaagagccc 300
acactgcgat gtgcatttta gtgcattaaa aggacttgca atgcactact gtctaatcca 360
gattacgacg acagccaatg gctttaagcg caagtatttt aaaactgaat tgtgagccgt 420
gattccagct ggagttaaat aggaaaactg ttgggcttaa agcagctgta gtggcaaata 480
ccaatcgagc atggcttcaa atcacgttca tttcgttcca catgacatga catcatggca 540
atggtatcaa agtgcacact tgtagacaca gaagcctcgg gtttgcttaa atttttatgg 600
cctcaacact taccaggtgc acagcttttc ttgccagcta ctggatgttc agtagcaggg 660
aaatcgacag ctggaagctc gagtatcaag aaaattgaga atgaatggaa aatgggttac 720
cactctcttt gttgcagaca aatttctcct tttaattgaa ggcaaaaagg gatatatatt 780
tttaaatagt tggcttctga cagagactta catgcgaaaa taaaaaattt agtattaatt 840
aaatcgactc gctgggaagt attgtcttat tttatattag aatttttaaa ttttaattta 900
gggacctata cataatggca atcaattagc ataaaatata acgttattaa tacattggtc 960
ctggaccttt gttcgtcggg gaatacgtat tttagtatat ctactgatgt gatctgctgt 1020
tgtagttggt ggatgaggtg gtcatcgact ggagctcagt ggtgcgcttt tggacatcca 1080
gctccatctt ggtggggcaa atcagaccca cccacagccg acggttcgtg gagtccaaga 1140
ggacaaagta aatgaacatc acaaatccct gcatagtggc cgtgatgcag aagaggtatg 1200
aaaaagccac acctgcctgc atgaaggcga atattccaaa gatccaggta agacccagca 1260
aaaagaagag cattatggag agtcggatct gcttaaccac cattttcttt tcgtttttat 1320
ggatgctctg gctcaaagag tgcgagatgc tgtagaaaac atacacgaat atcaccaagt 1380
tgcacacagt aatcaaggtc acagggagca ccaccccaaa aatgagacca tagcccgatg 1440
gatagcagat gccagtgtcc gtggaaagtt gggcagccga ggggacgtac gaatcgggat 1500
ctataagggc caccagcagg gtggggacca agggcagtag ccaggccaca atggcggcct 1560
tcaggatata gcgaggtggt ctctcgattc cgataactgt gacgtaacgc tggaactgca 1620
ggaaggcgat gattagcatc caactgaaga gaaccagaat ggagtactgc atggcagcgc 1680
ccagagccac acagcgaact gtgtttccat tcacaacaag cgcttcggat acatcatctg 1740
tattgagaaa cacgaagagc atcatctgga gacacatagc caggcaaagg tgcagtaaga 1800
ccttcgtgga ggcctgactg cgccagcttt tgaagagtgc ggcagtgagg aatattccca 1860
gtatacccaa aagggaaaga ctgcagccca ctatcgagat aatgtccagc actttctcat 1920
```

taatcggagt aattagaatc tcctcgccaa ggtcgttggc tcgatagctt ccacccacca 1980
ggaatgcaaa ctgtgtgagg tggttagtgt ggcactcgat gatggcatcc ttcaagaggt 2040
ccgaactgct ttcagtgctt actccctcag ttgaccaggt ttcatagttc caatatccgc 2100
atccactccc aatgctaaag gcctttgaat ccggattgcg aagattttca ttacgcagaa 2160
gaaagggcag tggcagagag agataattgc ctatttaatg gaagggtatt caaattagtt 2220
cgagtattcg aaaacaaatt ttcatgcgtt tcttacttct cagtcctggt attgaaatgg 2280
agattacttt gctgcgcggc ttacgagtgc gttgcaatga ggtttccacg aacagagcat 2340
cgtgggcata gactttaaag atgagatagg ttgctcccct gcttttcact tgccgccaga 2400
gattctccgg caaaaatgca gctgtctcca gattagattc ctccttgacc ttagccaagt 2460
cttcggagaa tcgaatgaag cggtaccaga agccactggc ggaagtcttt cgttgatccc 2520
ctggcgcaga gaatatagct aatccagtta tgcgatcgca ttcggggttg gcgaagaaca 2580
cgctgatatt gccggtaatc agagcttgga caccgatatc tgagaagttg tacgtttcca 2640
ctccagtagt ggccgtttct gcttcatcac tggtcggctt ggccaccacc ttgccgcatc 2700
tgtccttggg aaccagttgc tccggaagag catccatgta gctttcgaac tggctgagca 2760
gggaattggt agcgttgagc tgagctgata atctaagaac cttactgtcg caggacatta 2820
tctccttgca gacactgacc aagtccacac tgaccgtggc atccttgtcc tgctgcatca 2880
gggcgccgaa catttgtccg gtcatgtgaa catcggcggg tagaagattc ccacctctct 2940
gtcggaacat gttcctcatt atgccagtca tctcacgacg cccctgggtg tcattggtgc 3000
gcctcctgcc ctcgagaacg tcgtcatgca ggctattcaa gtcaccgctg atcgtgtgct 3060
ccctgaatcg cctcaaacac accacgggat cccagtgttc ggtggactcc cagttggctc 3120
ggttgtcctt cactctgcac tccctggtga ccggcattcc gtttcgcaga aggcaaacat 3180
cgtgcagggt ggcacgttcg ccgattctag cccgtttcca cttgttctgg tgaggtgttt 3240
gatcgtggtt acgcgttatg tagtcgtgct gaaactcctc cggctcgcag aacatggacg 3300
tacagtaaac cagatgtgtt tgcataagtt cgtccgatga attacgttcc tctccatcac 3360
agtactgaaa cgttccgttt ctcaagcaca tttgattttc ctttgggcag gcggaataaa 3420
tcggcgaaac caaggccaag accacaaaga gcccagcaca ccagttgaaa ttcatggcga 3480
tgaatcaggt gccacagctg tttgcagctc tccgaaaatg tttacaaatt ctttggcaga 3540
tagtactttt cacgtccgac gttcttgaaa tatcgcggca cactgagctg tctgtcgatc 3600
gattagattt ctgtagaccc cttcttcatg gtgtttgcta tcagttgcca ccacgagcgc 3660
taatcgaccc ccgttttttg tatatctaca cactactcac aaaaagacct caagtctgcg 3720
gcatttgata caatttacat tgtgattaac acgaccacgg caattgaagg cgaatccgac 3780
ggaacccttt gaggcaggga agtaacacca aacttgtgac gtcgtcttta tttgtgcaaa 3840
tagcgatttc acgccttttt accggtcttg gtacttaaaa tgctctggct gattcatttt 3900
ttacgcgctt gaattgccag ataagaagtt cttaccagga aacacggctg tacatgggct 3960
atttatacat caatgataca aaaataatgt taaaaaaaat ttaaaaactt taaaaactag 4020
ttttatttgc atccgttcaa gaaagagatt ttttttagat ttgtttgatc aattcattgc 4080
tatggcttct aataaatgtg aaatgtgaaa caccttcaca taaaaaagta tttacgcaat 4140
tcactcaagc aaaacagatc ctttgtcctt gaaaaaagtt atgaaatgca aagcacttca 4200
ccttggttgc ataaagtgcg attttcattt tgacaggcta cgcgttttca aagactcagc 4260
cttcgattat aatccacctg tactcacttc attgtaagcc agtttttggg ccaaaaagta 4320
atttttacag ccgacttcgc agcacgtaga ctgcgtgtgc tagcggaaat tgttagcttg 4380
ttgacacgtt t 4391

<210> 122
<211> 2325
<212> DNA
<213> Drosophila

<400> 122
atgtgcttga gaaacggaac gtttcagtac tgtgatggag aggaacgtaa ttcatcggac 60
gaacttatgc aaacacatct ggtttactgt acgtccatgt tctgcgagcc ggaggagttt 120
cagcacgact acataacgcg taaccacgat caaacacctc accagaacaa gtggaaacgg 180
gctagaatcg gcgaacgtgc caccctgcac gatgtttgcc ttctgcgaaa cggaatgccg 240
gtcaccaggg agtgcagagt gaaggacaac cgagccaact gggagtccac cgaacactgg 300
gatcccgtgg tgtgtttgag gcgattcagg gagcacacga tcagcggtga cttgaatagc 360
ctgcatgacg acgttctcga gggcaggagg cgcaccaatg acacccaggg gcgtcgtgag 420
atgactggca taatgaggaa catgttccga cagagaggtg ggaatcttct acccgccgat 480
gttcacatga ccggacaaat gttcggcgcc ctgatgcagc aggacaagga tgccacggtc 540
agtgtggact tggtcagtgt ctgcaaggag ataatgtcct gcgacagtaa ggttcttaga 600
ttatcagctc agctcaacgc taccaattcc ctgctcagcc agttcgaaag ctacatggat 660

```
gctcttccgg agcaactggt tcccaaggac agatgcggca aggtggtggc caagccgacc 720
agtgatgaag cagaaacggc cactactgga gtggaaacgt acaacttctc agatatcggt 780
gtccaagctc tgattaccgg caatatcagc gtgttcttcg ccaaccccga atgcgatcgc 840
ataactggat tagctatatt ctctgcgcca ggggatcaac gaaagacttc cgccagtggc 900
ttctggtacc gcttcattcg attctccgaa gacttggcta aggtcaagga ggaatctaat 960
ctggagacag ctgcattttt gccggagaat ctctggcggc aagtgaaaag caggggagca 1020
acctatctca tctttaaagt ctatgcccac gatgctctgt tcgtggaaac ctcattgcaa 1080
cgcactcgta agccgcgcag caaagtaatc tccatttcaa taccaggact gagaagcaat 1140
tatctctctc tgccactgcc ctttcttctg cgtaatgaaa atcttcgcaa tccggattca 1200
aaggccttta gcattgggag tggatgcgga tattggaact atgaaacctg gtcaactgag 1260
ggagtaagca ctgaaagcag ttcggacctc ttgaaggatg ccatcatcga gtgccacact 1320
aaccacctca cacagtttgc attcctggtg ggtggaagct atcgagccaa cgaccttggc 1380
gaggagattc taattactcc gattaatgag aaagtgctgg acattatctc gatagtgggc 1440
tgcagtcttt ccccttttggg tatactggga atattcctca ctgccgcact cttcaaaagc 1500
tggcgcagtc aggcctccac gaaggtctta ctgcaccttt gcctggctat gtgtctccag 1560
atgatgctct tcgtgtttct caatacagat gatgtatccg aagcgcttgt tgtgaatgga 1620
aacacagttc gctgtgtggc tctgggcgct gccatgcagt actccattct ggttctcttc 1680
agttggatgc taatcatcgc cttcctgcag ttccagcgtt acgtcacagt tatcggaatc 1740
gagagaccac ctcgctatat cctgaaggcc gccattgtgg cctggctact gcccttggtc 1800
cccaccctgc tggtggccct tatagatccc gattcgtacg tcccctcggc tgcccaactt 1860
tccacggaca ctggcatctg ctatccatcg ggctatggtc tcatttttgg ggtggtgctc 1920
cctgtgacct tgattactgt gtgcaacttg gtgatattcg tgtatgtttt ctacagcatc 1980
tcgcactctt tgagccagag catccataaa aacgaaaaga aaatggtggt taagcagatc 2040
cgactctcca taatgctctt ctttttgctg ggtcttacct ggatctttgg aatattcgcc 2100
ttcatgcagg caggtgtggc tttttcatac ctcttctgca tcacggccac tatgcaggga 2160
tttgtgatgt tcatttactt tgtcctcttg gactccacga accgtcggct gtgggtgggt 2220
ctgatttgcc ccaccaagat ggagctggat gtccaaaagc gcaccactga gctccagtcg 2280
atgaccacct catccaccaa ctacaacagc agatcacatc agtag                 2325
```

<210> 123
<211> 774
<212> PRT
<213> Drosophila

<400> 123

```
Met Cys Leu Arg Asn Gly Thr Phe Gln Tyr Cys Asp Gly Glu Glu Arg
 1              5                  10                  15
Asn Ser Ser Asp Glu Leu Met Gln Thr His Leu Val Tyr Cys Thr Ser
            20                  25                  30
Met Phe Cys Glu Pro Glu Glu Phe Gln His Asp Tyr Ile Thr Arg Asn
        35                  40                  45
His Asp Gln Thr Pro His Gln Asn Lys Trp Lys Arg Ala Arg Ile Gly
    50                  55                  60
Glu Arg Ala Thr Leu His Asp Val Cys Leu Leu Arg Asn Gly Met Pro
65                  70                  75                  80
Val Thr Arg Glu Cys Arg Val Lys Asp Asn Arg Ala Asn Trp Glu Ser
                85                  90                  95
Thr Glu His Trp Asp Pro Val Val Cys Leu Arg Arg Phe Arg Glu His
            100                 105                 110
Thr Ile Ser Gly Asp Leu Asn Ser Leu His Asp Asp Val Leu Glu Gly
        115                 120                 125
Arg Arg Arg Thr Asn Asp Thr Gln Gly Arg Arg Glu Met Thr Gly Ile
    130                 135                 140
Met Arg Asn Met Phe Arg Gln Arg Gly Gly Asn Leu Leu Pro Ala Asp
145                 150                 155                 160
Val His Met Thr Gly Gln Met Phe Gly Ala Leu Met Gln Gln Asp Lys
                165                 170                 175
Asp Ala Thr Val Ser Val Asp Leu Val Ser Val Cys Lys Glu Ile Met
            180                 185                 190
Ser Cys Asp Ser Lys Val Leu Arg Leu Ser Ala Gln Leu Asn Ala Thr
```

```
        195                 200                 205
Asn Ser Leu Leu Ser Gln Phe Glu Ser Tyr Met Asp Ala Leu Pro Glu
    210                 215                 220
Gln Leu Val Pro Lys Asp Arg Cys Gly Lys Val Val Ala Lys Pro Thr
225                 230                 235                 240
Ser Asp Glu Ala Glu Thr Ala Thr Thr Gly Val Glu Thr Tyr Asn Phe
                245                 250                 255
Ser Asp Ile Gly Val Gln Ala Leu Ile Thr Gly Asn Ile Ser Val Phe
            260                 265                 270
Phe Ala Asn Pro Glu Cys Asp Arg Ile Thr Gly Leu Ala Ile Phe Ser
        275                 280                 285
Ala Pro Gly Asp Gln Arg Lys Thr Ser Ala Ser Gly Phe Trp Tyr Arg
    290                 295                 300
Phe Ile Arg Phe Ser Glu Asp Leu Ala Lys Val Lys Glu Glu Ser Asn
305                 310                 315                 320
Leu Glu Thr Ala Ala Phe Leu Pro Glu Asn Leu Trp Arg Gln Val Lys
                325                 330                 335
Ser Arg Gly Ala Thr Tyr Leu Ile Phe Lys Val Tyr Ala His Asp Ala
            340                 345                 350
Leu Phe Val Glu Thr Ser Leu Gln Arg Thr Arg Lys Pro Arg Ser Lys
        355                 360                 365
Val Ile Ser Ile Ser Ile Pro Gly Leu Arg Ser Asn Tyr Leu Ser Leu
    370                 375                 380
Pro Leu Pro Phe Leu Leu Arg Asn Glu Asn Leu Arg Asn Pro Asp Ser
385                 390                 395                 400
Lys Ala Phe Ser Ile Gly Ser Gly Cys Gly Tyr Trp Asn Tyr Glu Thr
                405                 410                 415
Trp Ser Thr Glu Gly Val Ser Thr Glu Ser Ser Ser Asp Leu Leu Lys
            420                 425                 430
Asp Ala Ile Ile Glu Cys His Thr Asn His Leu Thr Gln Phe Ala Phe
        435                 440                 445
Leu Val Gly Gly Ser Tyr Arg Ala Asn Asp Leu Gly Glu Glu Ile Leu
    450                 455                 460
Ile Thr Pro Ile Asn Glu Lys Val Leu Asp Ile Ile Ser Ile Val Gly
465                 470                 475                 480
Cys Ser Leu Ser Leu Leu Gly Ile Leu Gly Ile Phe Leu Thr Ala Ala
                485                 490                 495
Leu Phe Lys Ser Trp Arg Ser Gln Ala Ser Thr Lys Val Leu Leu His
            500                 505                 510
Leu Cys Leu Ala Met Cys Leu Gln Met Met Leu Phe Val Phe Leu Asn
        515                 520                 525
Thr Asp Asp Val Ser Glu Ala Leu Val Val Asn Gly Asn Thr Val Arg
    530                 535                 540
Cys Val Ala Leu Gly Ala Ala Met Gln Tyr Ser Ile Leu Val Leu Phe
545                 550                 555                 560
Ser Trp Met Leu Ile Ile Ala Phe Leu Gln Phe Gln Arg Tyr Val Thr
                565                 570                 575
Val Ile Gly Ile Glu Arg Pro Pro Arg Tyr Ile Leu Lys Ala Ala Ile
            580                 585                 590
Val Ala Trp Leu Leu Pro Leu Val Pro Thr Leu Leu Val Ala Leu Ile
        595                 600                 605
Asp Pro Asp Ser Tyr Val Pro Ser Ala Ala Gln Leu Ser Thr Asp Thr
    610                 615                 620
Gly Ile Cys Tyr Pro Ser Gly Tyr Gly Leu Ile Phe Gly Val Val Leu
625                 630                 635                 640
Pro Val Thr Leu Ile Thr Val Cys Asn Leu Val Ile Phe Val Tyr Val
                645                 650                 655
Phe Tyr Ser Ile Ser His Ser Leu Ser Gln Ser Ile His Lys Asn Glu
            660                 665                 670
Lys Lys Met Val Val Lys Gln Ile Arg Leu Ser Ile Met Leu Phe Phe
```

```
        675                 680                 685
Leu Leu Gly Leu Thr Trp Ile Phe Gly Ile Phe Ala Phe Met Gln Ala
    690                 695                 700
Gly Val Ala Phe Ser Tyr Leu Phe Cys Ile Thr Ala Thr Met Gln Gly
705                 710                 715                 720
Phe Val Met Phe Ile Tyr Phe Val Leu Leu Asp Ser Thr Asn Arg Arg
                725                 730                 735
Leu Trp Val Gly Leu Ile Cys Pro Thr Lys Met Glu Leu Asp Val Gln
            740                 745                 750
Lys Arg Thr Thr Glu Leu Gln Ser Met Thr Thr Ser Ser Thr Asn Tyr
        755                 760                 765
Asn Ser Arg Ser His Gln
    770


<210> 124
<211> 7986
<212> DNA
<213> Drosophila

<400> 124
gtgttaatcc ccttggctta tcggtttaat ctatttataa aatcaattca aacataaagc 60
ggcgcatatc tcagcgatgt tgggtaggta accaaaagta aatacactct aattcctatt 120
gatcttgtat aaatacgagt ggatagttag ttagatatca ccattttgca ccgtgcataa 180
acataaagtg agatactcgt tttcgagaga tgcacttctg agtacccgta ttcaaatcgt 240
actctcagta gccaagtact cagtactcaa gctgtttgtc ttacgggtgt ggacgtgcga 300
gcagacactt taggagtggg atagccggag cagaccggct gaagtgtcat ccgatcgaag 360
gtcgcctcaa cagcacagcg atttcgattc cgattccaaa tccgattctt tggccatagg 420
aattcctctg ggaggtggag catgtggctg atttatttat ttcgagcgat ataaacagag 480
caatacgaga atgggcgaat ataagtttgc tcgaacggct ggtccgttcc gagctgaaat 540
tgaattgtaa atggccatat taattgggaa tcatttaaaa tttccaatga ttttttctgt 600
gttttggtgt cggcagtcta tgcatgttcg gtggcataat ttatgggcgc taattactat 660
cgccccattt tgtgtatttc gcgtgctgtc caaggccaat tacgcttact tattgcaatt 720
tatgtggcca cagtccggcc gtaattggaa attgtgtaag gtcgcaggca atttcattat 780
ttgtatgcac ttctggacgc tccggttgcc gggtctagca tttttctagc atttttctaa 840
tgacgccagc gcgtttgggg acgacgttac ctagcccagt tttccctttg cagccatttg 900
attaagtgac ctacgtttcg cccacacaga atcatcaatt tgttcatgtc gccgacggtg 960
cgaataaggt ccaacctaag tgagtgagct atccgccggc atggcggcgt tggaacatcc 1020
ggatgccatg ctgccatact ccaaattgtt gtcaacagcc ctcactgtgg ccagtggtaa 1080
ctcgaactcc tccgatggta ctttgctgct catccagacc agtgcggcac caggtggcga 1140
tgtgatggcc accggagccg gagccggaaa gttgggtgtg ccctcgatga actcgtacct 1200
ggtgagcatg gcgtggccca agtcgctggc agtggccgtc ttcctggtca tcatcctggt 1260
caccgtggtg ggcaacacac tcgtcatcct cgccgttctg accacacgcc gcctgcgcac 1320
cgtaaccaac tgtttcgtga tgaacttggc catcaccgac tggctggtcg gcacctgcgt 1380
gatgcctcca tcggtggtgc tctacataac aggtgagaat ggcagcggag ggcctgtcca 1440
tgcccatcga tggccaaaga tttatggctc tgtgggggct ttacacttcc ggggaaggtg 1500
gaatcggggc aagtgtttac tattcgaagc cctccggtgg aagttcgggt agtaaacatt 1560
tgcccaaatc gatgggggag cacatttggt cgatttcgat ttaaaataat cccacagcga 1620
aaattataaa atttctccgt tggcaaaata tacatatttt tgattgcaaa atatcttgtg 1680
agtggtgtgc gagtgtttta tgggtttcgc tggaaaacct ttgataattt tttctggaat 1740
ctatattttt agcagaattt tgaataaagc ataacaacat tagtggttaa aaattttgta 1800
taatttagta attttaaata aattatgtat ctttaataat taacatctct tatttaacgc 1860
atttcataaa tgtataaaaa gcaatcaaac atgtgatctt ttcgaagcca aatgaatata 1920
caaagtttaa tattccttgt cataacttga acaatacaaa gaagaattta tagtggttta 1980
aattgaagga tctgccagtg taatgcctcc ttaatcccgg ccaggtgcac aaaccgattg 2040
aagcacggcg acaaatattt gttctttgca ggcacttggc gctttggctg gatactgtgc 2100
gacatttgga tctcgctgga catcctgctc tgcagcggct ccatcctcag cctgtgcgcc 2160
atcagtctgg acaggtgagt tggccaacta atttcgtaat ttgccgttga tgtcatagca 2220
atgaccacag tccatagccc gttgtccacg gtcctcggtc cttagataaa gtcatggcaa 2280
ggtgcaaaaa cagtgcaaat aatttatttg ttgacgtttt tattgccgca tgaccgttgt 2340
```

```
cctggccatt cccagtgcat atcctttgag tgtgcgacgc cacaatggca aggattgggt 2400
tgggcttgga ttcccattca cccattccct atttccccat ttcctccatt tagccatccc 2460
accgtttgct tttgttggaa gtttcaatta cattttaatg gcatttccaa gttttgtgtt 2520
gtggctcccg agttgagctg atccagctga gccatggagc catggacctg aacctcaacc 2580
aggacaaaca ccgacgccca ctggcgttgt gtgtgccttt tggcctggcc agaaagccgt 2640
ctgcggcttt gcgaacaata tgtccgcttg gcttttgttt ttttccagca tttctctggc 2700
catttccgtt gcgtttattg gcccggcttt tataatttcc tctttttcac cttcgttttt 2760
gttttatatt tcgcaggtat ctcgccgtca cacaaccact gacgtactcc aagaagcgcc 2820
gatccaagcg actggccctc ctcatgattc tcgtcgtgtg gataactgcc ctgtcaatca 2880
catgtccacc ctatttgggc tggtaagtta gcgccagctg cacggagaac aaaatcaacc 2940
ttaacccatg atggatcgcc tattcatctg ttaagaaaac cattcatgtt ataatcattt 3000
aattgaacgc attaaattgg gttgggagta caggcaacat gtgataatga tcagtatcca 3060
aaggaagaaa aataatatat ttaatatcat ataatattat taataatatt atattataaa 3120
tatggccaag ataaatttat ttgtagtgct ctattaagtt tatagatata tattttttttg 3180
aattttgtaa tattattttc actctttccc gtgtaattcg aaatattata tgtaatattt 3240
ttattgagca cccatttttt ttttaccaaa tttgtccttc cggaggcgtg acaatacttt 3300
gtggcttttt gtggctgaaa gtcctgctcg cttttttttt tcttttttga gctaatggct 3360
gtgcggtggc gcagcataga taatgtgcta caaaagggtc tcgaaataca cacgtttatt 3420
tatattatat ttgcgccttt ggctgcgcct aaaaaacatt acgtaattga cgacgggttg 3480
ttgctgttgc tgctggtcgt gcagccatta cattgaagtt ttaatgagtt gttaatgact 3540
ttgtttgccg cggcgcattt gctttgttct cacctacgca cacacactca gcgcgaaagc 3600
cgcacacttg tgtgttgcat acttccaggc gcccaaatta accagcactc actcacacac 3660
actcgcactc actcacacac accacaaagc cagcatgtcc ttaatgccat tggccatggc 3720
ttagttttgg ccccctttgg ctgaaagcga agagcatggc ttaaagggca cattaaaacg 3780
gtctgaggcc gaattacagc cacagctgca ggacattgga aaatttggaa aatttggcgc 3840
tgggggaaaa tcgcaattat aagcgccagc agatggcgcc acttacatat tttcattctt 3900
taagcggact agttttatga aaaaagtttg caaatgcagc ggttgcaaca actgtttgtt 3960
ttactcacaa tctgttattt gttattgcct aatgaatact aaactgattg tgtttgcttg 4020
ctatttaaat ttagatttaa cttttggcta attgatttaa gcattgttta tttagttcat 4080
tttttaattt aactaatgca tagctaatgt attaaattgc cctcaatttc ggattgaata 4140
attaatgtta aacaagaagt ttagtaggaa aatcggctat ttccattcgg gtcttgtaaa 4200
taccaacctt aacaaagttg catttgtaca attatactgt caggttattt tggtatcggg 4260
tagcagtaat taccgtgagt ttttggataa cacatacttg gggtaacaag cagctagttt 4320
agttgaatta caacaccgga aaattccatt gacagtagtg acaatgaggc aataaacttt 4380
ttgaaagaga aacattatag gaaactgtgt atccagtggg cgtgccccaa actttggttg 4440
gcatatttgt aaaaattggc catccaaaaa ataaatcaaa acaactggca atcattttat 4500
aatcgtgtgg gcgtgagtct tttaggcgat tggggggctt tagagtgggc gtggtaatgt 4560
tctgaaataa gcttgttccc gagatttgga ctccgggggga caaatggaca taaccgaatc 4620
aaggatatat atacttcatg cttctgtctg ctacatgata ttcagtgaac ctagtattct 4680
aagaatagtg ggtacaataa cccagtatga actggaagaa gctgaatctt gccacctaca 4740
gtggatgtat gtgacaattt gcccattttt tttactcaat tccccacagc acacgggtgt 4800
acctttagtt gacacatttg gtgtacgcca ccacattact ctgttgtctc ggaatcgttt 4860
tcttccgcat attgatttaa agtcagctcc gatttgtact gcctgaatga ccaacacact 4920
gcgatgtgca taagtaaaag gcttacaggt gggactactc tgactggctt aggtttagga 4980
tttgaattcg aatttcgccc ggcttaactc agccagttgt cattccccaa gtcccagttc 5040
cccagccgca atcacaaccg caagagccac caaaccagag caggtaccag gtacctccca 5100
catccaccac atctcacatc cttccaaccc gcaaaccgca tgcaattcga actggaattc 5160
aatttgccat tcacatccac attcgcaggt acgaggtcgg tcggcaccag gcggagttcg 5220
tggactgccg ctacaaccaa aacaagggct acgtggtctt ctcggcgatg ggatcgttct 5280
tcattcccct caccgtgatg ctctacgttt acgtgaagat tggttatgtc ctcacgtcac 5340
ggcgacagcg cattgtgcgc gatgcggtaa gctacatctc cagctccatc actatataag 5400
catatatgta tgtgtccacc actttggtgt cccctgaaat gtgtccggaa atttatcatg 5460
tcacacctaa cacatctaac cagctaacca ggggactaag ttgtcaggca ctggggaaaa 5520
agatccagga gtgcgctcaa cacttcaagt gcagatcttc attaagttgg gatacccact 5580
gggtatccca cttgagtacc agccacttat cgcacagtac tgcaggtgat attttaagtg 5640
ctgagctttg ataacagaag tccaactaat atatctaatc acaaattaaa tttgctgtga 5700
attcttggat gtaaactgaa tgaattatct tttgtttgtt agcatgaaaa atgaacttca 5760
gtgtcaataa atgaaatgaa aatgcaataa gaacactttt tgaatattta aaaagagtgc 5820
atagtagttg ataaatatat tcttttagtt tagtactaat tgaacgtttg ctgccaagag 5880
caagtgaagt aaattgtaga aagtaggtga cgtgaaatgg aacgaagtaa ttacagcaat 5940
```

ccaaatcccg aatctctatg ccctgacatg caatctaaat tattcctcac gtttttgtca 6000
gtgctctgct tacatattcc tcttctatta tgtccttcca acagtactcg gagcgcacgg 6060
cagactacga tgtggatggg gacaacttca tctcggagtc ggagcactat cattgcacgc 6120
caaccaagtg gctgccgaac aggaagtccc gctggaggtt caactccctg cacgacccac 6180
cacccagcgc cgccacgaat gcagcagcga agggacacca gtcgtccgtc aagtgctcca 6240
agtgctccaa gtgctcctct gccactggaa ctgctggagc tggagctgga gctggtgtgg 6300
tagacaaaac gctgacgttc aaacatcagc ccacctttta cgagctggtc gaggtgtcgc 6360
gtctctcgtc gctcattcac tgctcggcaa ttagctgcaa gtatggagga ccctgcatgc 6420
actccacgcc ctcgggcatg cactacgttg gcacggaggc ctccttctcg gacacctgcc 6480
tgggtggcaa tgtgagcaac aacaccacca ccaccaccac caacaccgtc gcactggaaa 6540
ccaacgagaa actggaggcg gaggccaagc aaaagcaaca gcagcaacga tcctccttcg 6600
accaccatca cccgcatgca caccatcatg gctcccatgg ccagcagagc catcacaatc 6660
atcaccatcg catgccgatg cgagtttcga ccacgaaacg tgatagcaag accgccaaga 6720
ccctgactat tgtgatgggc ggcctaattg cttgctggct gcccttcttc gtctattatc 6780
tgctgatacc cttcctgccg cgacccgccg tcctcgagga cctcatgttc ggcttcacct 6840
ggattggctg ggtcaactgc gccatcaacc cgttcatcta cgccttctac aatccggact 6900
ttcgcactgc cttctggcgg ctcacctgtc gccccatttg caagcagaag cgtccgccca 6960
accacctggc catgttccgt ggctagtttc cggatgcaga tgcagttgcg gatgcggatg 7020
cggatccgga accggtttcc gaatcactga gcacttgagt gtgtgtcata gtgccgccat 7080
ctccagtcgc acaatcgacg ctcacttggg tggtggggga aatatctcag gcatatcctt 7140
aacttttcgc cgtgctggtt caccaacttc tatgtaatta gcataaaaaa taagtggaga 7200
aactaccaac tcgcatttcg atctttttga agtacaagtt atgccaatgc aaggacatca 7260
aattgactct agactacagt acagatagta aataaaaaag aatttaagga gtattaacta 7320
caatagataa gctgtttgag aaatgcacta tctcttatat taaatcaatt gaagataatg 7380
aaaatcaaat attagcttaa caaaactgct gaatttaatt caagtttatt tcatgaaaac 7440
aaatatatag aaactagttt aaaatttatc acacatgtcc ctccaaaaat tgtaactaat 7500
attttggcaa acgaagccaa tggagggacc ttttggccaa ctcgtcgcca ctagctcacc 7560
aactaagtat tcatttgcca ggcactataa gctgtttgcc gcactctatc tgcatcgttt 7620
atgccaaacg cagctagtaa aactaagata cttatggtca atattgattg gagcaaagat 7680
tgccttgtta aacccaaaaa atcgactgcc ttttgaagct tcaaaagccg catttcgagc 7740
agcatcttaa aggtcacttt acggtcgcaa tttttttttt ttaagactca ctggacttgg 7800
ccggaagttg agattcttgc ctagaccaga tccaaatcct atatatagta ttttcctgtg 7860
tgtgagcgtt gattgtgcga ggcgagcggc acgaagcgaa acgtaacgaa acaaaagaca 7920
aattatagta tttgcgagtc tttgttgcct gacggcaact gttggccatc gggatgagtg 7980
ggatga 7986

<210> 125
<211> 1704
<212> DNA
<213> Drosophila

<400> 125
atggcggcgt tggaacatcc ggatgccatg ctgccatact ccaaattgtt gtcaacagcc 60
ctcactgtgg ccagtggtaa ctcgaactcc tccgatggta ctttgctgct catccagacc 120
agtgcggcac caggtggcga tgtgatggcc accggagccg gagccggaaa gttgggtgtg 180
ccctcgatga actcgtacct ggtgagcatg gcgtggccca agtcgctggc agtggccgtc 240
ttcctggtca tcatcctggt caccgtggtg ggcaacacac tcgtcatcct cgccgttctg 300
accacacgcc gcctgcgcac cgtaaccaac tgtttcgtga tgaacttggc catcaccgac 360
tggctggtcg gcacctgcgt gatgcctcca tcggtggtgc tctacataac aggcacttgg 420
cgctttggct ggatactgtg cgacatttgg atctcgctgg acatcctgct ctgcagcggc 480
tccatcctca gcctgtgcgc catcagtctg gacaggtatc tcgccgtcac acaaccactg 540
acgtactcca agaagcgccg atccaagcga ctggccctcc tcatgattct cgtcgtgtgg 600
ataactgccc tgtcaatcac atgtccaccc tatttgggct ggtacgaggt cggtcggcac 660
caggcggagt tcgtggactg ccgctacaac caaaacaagg gctacgtggt cttctcggcg 720
atgggatcgt tcttcattcc cctcaccgtg atgctctacg tttacgtgaa gattggttat 780
gtcctcacgt cacggcgaca gcgcattgtg cgcgatgcgt actcggagcg cacggcagac 840
tacgatgtgg atggggacaa cttcatctcg gagtcggagc actatcattg cacgccaacc 900
aagtggctgc cgaacaggaa gtcccgctgg aggttcaact ccctgcacga cccaccaccc 960

```
agcgccgcca cgaatgcagc agcgaaggga caccagtcgt ccgtcaagtg ctccaagtgc 1020
tccaagtgct cctctgccac tggaactgct ggagctggag ctggagctgg tgtggtagac 1080
aaaacgctga cgttcaaaca tcagcccacc ttttacgagc tggtcgaggt gtcgcgtctc 1140
tcgtcgctca ttcactgctc ggcaattagc tgcaagtatg gaggaccctg catgcactcc 1200
acgccctcgg gcatgcacta cgttggcacg gaggcctcct tctcggacac ctgcctgggt 1260
ggcaataaac tggaggcgga ggccaagcaa aagcaacagc agcaacgatc ctccttcgac 1320
caccatcacc cgcatgcaca ccatcatggc tcccatggcc agcagagcca tcacaatcat 1380
caccatcgca tgccgatgcg agtttcgacc acgaaacgtg atagcaagac cgccaagacc 1440
ctgactattg tgatgggcgg cctaattgct tgctggctgc ccttcttcgt ctattatctg 1500
ctgataccct tcctgccgcg acccgccgtc ctcgaggacc tcatgttcgg cttcacctgg 1560
attggctggg tcaactgcgc catcaacccg ttcatctacg ccttctacaa tccggacttt 1620
cgcactgcct tctggcggct cacctgtcgc cccatttgca agcagaagcg tccgcccaac 1680
cacctggcca tgttccgtgg ctag                                        1704
```

<210> 126
<211> 567
<212> PRT
<213> Drosophila

<400> 126

```
Met Ala Ala Leu Glu His Pro Asp Ala Met Leu Pro Tyr Ser Lys Leu
 1               5                   10                  15
Leu Ser Thr Ala Leu Thr Val Ala Ser Gly Asn Ser Asn Ser Ser Asp
            20                  25                  30
Gly Thr Leu Leu Leu Ile Gln Thr Ser Ala Ala Pro Gly Gly Asp Val
        35                  40                  45
Met Ala Thr Gly Ala Gly Ala Gly Lys Leu Gly Val Pro Ser Met Asn
    50                  55                  60
Ser Tyr Leu Val Ser Met Ala Trp Pro Lys Ser Leu Ala Val Ala Val
65                  70                  75                  80
Phe Leu Val Ile Ile Leu Val Thr Val Val Gly Asn Thr Leu Val Ile
                85                  90                  95
Leu Ala Val Leu Thr Thr Arg Arg Leu Arg Thr Val Thr Asn Cys Phe
            100                 105                 110
Val Met Asn Leu Ala Ile Thr Asp Trp Leu Val Gly Thr Cys Val Met
        115                 120                 125
Pro Pro Ser Val Val Leu Tyr Ile Thr Gly Thr Trp Arg Phe Gly Trp
    130                 135                 140
Ile Leu Cys Asp Ile Trp Ile Ser Leu Asp Ile Leu Leu Cys Ser Gly
145                 150                 155                 160
Ser Ile Leu Ser Leu Cys Ala Ile Ser Leu Asp Arg Tyr Leu Ala Val
                165                 170                 175
Thr Gln Pro Leu Thr Tyr Ser Lys Lys Arg Arg Ser Lys Arg Leu Ala
            180                 185                 190
Leu Leu Met Ile Leu Val Val Trp Ile Thr Ala Leu Ser Ile Thr Cys
        195                 200                 205
Pro Pro Tyr Leu Gly Trp Tyr Glu Val Gly Arg His Gln Ala Glu Phe
    210                 215                 220
Val Asp Cys Arg Tyr Asn Gln Asn Lys Gly Tyr Val Val Phe Ser Ala
225                 230                 235                 240
Met Gly Ser Phe Phe Ile Pro Leu Thr Val Met Leu Tyr Val Tyr Val
                245                 250                 255
Lys Ile Gly Tyr Val Leu Thr Ser Arg Arg Gln Arg Ile Val Arg Asp
            260                 265                 270
Ala Tyr Ser Glu Arg Thr Ala Asp Tyr Asp Val Asp Gly Asp Asn Phe
        275                 280                 285
Ile Ser Glu Ser Glu His Tyr His Cys Thr Pro Thr Lys Trp Leu Pro
    290                 295                 300
Asn Arg Lys Ser Arg Trp Arg Phe Asn Ser Leu His Asp Pro Pro Pro
305                 310                 315                 320
```

Ser Ala Ala Thr Asn Ala Ala Ala Lys Gly His Gln Ser Ser Val Lys
325 330 335
Cys Ser Lys Cys Ser Lys Cys Ser Ser Ala Thr Gly Thr Ala Gly Ala
340 345 350
Gly Ala Gly Ala Gly Val Val Asp Lys Thr Leu Thr Phe Lys His Gln
355 360 365
Pro Thr Phe Tyr Glu Leu Val Glu Val Ser Arg Leu Ser Ser Leu Ile
370 375 380
His Cys Ser Ala Ile Ser Cys Lys Tyr Gly Gly Pro Cys Met His Ser
385 390 395 400
Thr Pro Ser Gly Met His Tyr Val Gly Thr Glu Ala Ser Phe Ser Asp
405 410 415
Thr Cys Leu Gly Gly Asn Lys Leu Glu Ala Glu Ala Lys Gln Lys Gln
420 425 430
Gln Gln Gln Arg Ser Ser Phe Asp His His His Pro His Ala His His
435 440 445
His Gly Ser His Gly Gln Gln Ser His His Asn His His His Arg Met
450 455 460
Pro Met Arg Val Ser Thr Thr Lys Arg Asp Ser Lys Thr Ala Lys Thr
465 470 475 480
Leu Thr Ile Val Met Gly Gly Leu Ile Ala Cys Trp Leu Pro Phe Phe
485 490 495
Val Tyr Tyr Leu Leu Ile Pro Phe Leu Pro Arg Pro Ala Val Leu Glu
500 505 510
Asp Leu Met Phe Gly Phe Thr Trp Ile Gly Trp Val Asn Cys Ala Ile
515 520 525
Asn Pro Phe Ile Tyr Ala Phe Tyr Asn Pro Asp Phe Arg Thr Ala Phe
530 535 540
Trp Arg Leu Thr Cys Arg Pro Ile Cys Lys Gln Lys Arg Pro Pro Asn
545 550 555 560
His Leu Ala Met Phe Arg Gly
565

<210> 127
<211> 3681
<212> DNA
<213> Drosophila

<400> 127
tccatgcatt tataattcaa cattcgctcg cggccttatg aaaagctcat ttggttacga 60
gtgtcctggg tccttaagtc ctttctacct cgaaagtgtt aatgagccgg caggctgccg 120
taattatgca aatttttaat gactatgaaa atggtttccg cggagcaaca agtgctcgcg 180
tcttggtcct tggatctcta taccсctcat acccctctaa aaatgccagc taattccatt 240
tgaccccaga ggccccaaag caatccattt gcatggctac caatagctcg tcgcattcgt 300
cgtaccaacc caaccatttt ccatttaatc cctacgatgt gagagctctg tactgagctt 360
actgctggca ttcggtcttt agatttatac agattttgtc aaacaaaaca tgaattcgca 420
cttttacaca catatgtgtg tctgtaaatg actgaaaggg agagcgtcga aagccgattt 480
ggtgccattc agttgaagcc ctttaggtcg tgattcgcat atggtcgagg ccatttaagc 540
tgatatcagg cacattaaat gtagtaaagg aacgtaagta tgataaaggc agcatcacaa 600
aagtgtcaag taaatgcaaa ttaacattaa tttaagactt gaattcgtct gtactgaaac 660
agtacaatta aaatataatt agctatatat ttgtaacttt taagaacgga tacttctggg 720
accaatgttt tttgaacata acagatatct accatatgaa gtgaaataaa tagaattttt 780
cgggaccatt ttgtttccag atttcaaatg aaatgaaacg catatttttg ccagtttttg 840
ataataaata tcaggggact tggccaaatt ttaatacaat ctcatttaga aaaataaaac 900
aattgcatgc cgataaaata ataagtgata ccgaaagaga gcttttgatt agagtgaaat 960
atgctcatgc ggttcagtcg agatccaaag tttgcgcaga atgcttctga atattttagc 1020
tataattctt gtattcgtaa tttcaagcca atcggaggca gttattcctg gatgtgatta 1080
cttcgacacc gtggatatat ctcatatccc aaagctaaat gactcctatg cgtacgagga 1140
gctaatcatt cctgctcatc tcacgggtct atatacattt aggcaactgg cggatggatc 1200

acaggagcca gtgaagagtc acttaagagc atgtatctgc aaactaaaac cttgtatccg 1260
attctgctgc cccgcaaca aaatgatgcc gaatagtcgc tgtagcgatg gactcacaga 1320
gaatctgaag cggatcaatc cctacttgaa gatcacactg gaggatggaa ctataggaaa 1380
atattacctt ctcaccgaca tgatcgtctt aagatatgag tttcggtatt gcgagaaggt 1440
tgtttcggta caggaggatc agtacaagct atatgaggtt aagatttcaa tctttttgaa 1500
gtatctcagt ataatttgat atctatagaa tggcagcttt atgattaaac ccgatgtaaa 1560
ctggacattg agcaagcagt ggtactgcct ccatccccgt ttggaagatc cgaattccat 1620
atggatactg gaacatgtat atataccgaa gtcaatgcca gccgtacctc agggcaagca 1680
cctactatat catctaatct catctcatct aattctctga tttttccagt cggtactatt 1740
tccatggttg gttgtattct cacgattgca gtgtatctct acataaagaa actgcgaaat 1800
ttgcttggaa aatgtttcat ttgctatgta ttttgtaagt ttgtgcagta ccttatctgg 1860
gcaggaggag atttgaactt gtggaataat atttgctccc tagcaggtta gttgaatcca 1920
tattaaagtt tgacttaaaa tatgcgttta ttctaaacag gttacactaa ctatttcttt 1980
gctttggctt ctcacttttg gctctctgtt atgagtcatc aaatatggaa aaacttgagg 2040
ttaatcaatc gggatgaacg tagctatcac tttctcatct acaacatcta cggttggggc 2100
acacctgcta tcatgacggc aattacatat ctggtggatt gggcctggga ggatagacct 2160
gacaaactga attggattcc cggtgttggc ttatatcgat gctggataaa cagtgggttt 2220
tagggtagag tgtttaagcc agtttaagac cagtaattga ttttctttga acagcctatg 2280
actggtctgc tatgatatac ctatatggac cgatgctgat cttgagttta ttcaatgtgg 2340
tcacgtttat cctgacagta aatcacataa tgaagattaa gagtagcgtt aagagttcta 2400
cccaacagca gcgaaagtgt atacagaata atgagtaagc tttaagtatt tataacttac 2460
tatttctaat tccattctct tactatagtt ttctattata cctacgactg tcggtgatga 2520
tgggcgtgac tgggatttct gaggtaataa cttacttcgt caagcggcac aaattttggc 2580
gacaagtcct tcgggtgcca aatttttttcc atttgggttc cggtatagtc gtatttgtgc 2640
tatttattct taagcgcagc acatttcaaa tgataatgga gaggtaagag ctttttgctt 2700
aatagatttt aaagttacta tatcaatatc attttagaat cagcggtccg aggagacaac 2760
aacctgcatc ttgatcaacc tactagttta aataaatgca aattatagat aagtgagctc 2820
tttttttta taagtttagt atagagcagt ccctctatgg tcaataaatg cttataatct 2880
tcttattatt gaaaaattct ttattattga tatttaatca ataatatatt tttatttttt 2940
tttttttgt tagaacaaaa tgtcaattgt catcaaaaat atgtgtttgt gaataaaatc 3000
aagttccgta ggaaaatatg cacttttatg gaaactcgct ttttcaatta gtccttgcat 3060
acgcgccagg accccaaatc cttccacctt ccccacaatc ctcggctggt gtcagaatgc 3120
aaatgcagat ttttaagcct tgcggcagct tttgcgcaaa cagattgtaa acaaataaat 3180
gccgccggca gacacacaga tactcgaaca ttcggtggtc agtggccaga ttgcagcctg 3240
aaattcattc aatactttat ctcacatgca ttgcacttca cttttgagcc acgttaagtt 3300
cagtctgccg aaaccttttt cgcacttaag tggttaaatt ggctgtgata aattgttgcc 3360
gagtcaaagt tgctgccgct gcggggcacc tcaagtcaca tgtggtgatg gtgcgagtgg 3420
gtggttgggt ggcgtgtgct ttctcctcg gaaaatgcgt gaaaagtcca caagcacaaa 3480
aagtctggga ggctggctct ctctcatctc atcacgttca ttgagtgcct gtcgctgaca 3540
gtaatcataa cgcagtcatc agctcgtaat attacggccg cctgacatat gacatctggg 3600
aaagcggaaa atgggggctc agggaccaag gaaaaggtat gcgtactatt aatatgccag 3660
tgtgtgaagg agctttcatt c 3681

<210> 128
<211> 1503
<212> DNA
<213> Drosophila

<400> 128
atgcttctga atattttagc tataattctt gtattcgtaa tttcaagcca atcggaggca 60
gttattcctg gatgtgatta cttcgacacc gtggatatat ctcatatccc aaagctaaat 120
gactcctatg cgtacgagga gctaatcatt cctgctcatc tcacgggtct atatacattt 180
aggcaactgg cggatggatc acaggagcca gtgaagagtc acttaagagc atgtatctgc 240
aaactaaaac cttgtatccg attctgctgc cccgcaaca aaatgatgcc gaatagtcgc 300
tgtagcgatg gactcacaga gaatctgaag cggatcaatc cctacttgaa gatcacactg 360
gaggatggaa ctataggaaa atattacctt ctcaccgaca tgatcgtctt aagatatgag 420
tttcggtatt gcgagaaggt tgtttcggta caggaggatc agtacaagct atatgaggtt 480
aagatttcaa tctttttgaa gtatctcagt ataatttgat atctatagaa tggcagcttt 540
atgattaaac ccgatgtaaa ctggacattg agcaagcagt ggtactgcct ccatccccgt 600
ttggaagatc cgaattccat atggatactg gaacattcca tggttggttg tattctcacg 660

```
attgcagtgt atctctacat aaagaaactg cgaaatttgc ttggaaaatg tttcatttgc 720
tatgtatttt gtaagtttgt gcagtacctt atctgggcag gaggagattt gaacttgtgg 780
aataatattt gctccctagc aggttagttg aatccatatt aaagtttgac ttaaaatatg 840
cgtttattct aaacaggtta cactaactat ttctttgctt tggcttctca cttttggctc 900
tctgttatga gtcatcaaat atggaaaaac ttgaggttaa tcaatcggga tgaacgtagc 960
tatcactttc tcatctacaa catctacggt tggggcacac ctgctatcat gacggcaatt 1020
acatatctgg tggattgggc ctgggaggat agacctgaca aactgaattg gattcccggt 1080
gttggcttat atcgatgctg gatatggtct gctatgatat acctatatgg accgatgctg 1140
atcttgagtt tattcaatgt ggtcacgttt atcctgacag taaatcacat aatgaagatt 1200
aagagtagcg ttaagagttc tacccaacag cagcgaaagt gtatacagaa taatgagtaa 1260
gctttaagta tttataactt actatttcta attccattct cttactatag ttttctatta 1320
tacctacgac tgtcggtgat gatgggcgtg actgggattt ctgaggtaat aacttacttc 1380
gtcaagcggc acaaattttg gcgacaagtc cttcgggtgc caaatttttt ccatttgggt 1440
tccggtatag tcgtatttgt gctatttatt cttaagcgca gcacatttca aatgataatg 1500
gag                                                               1503
```

<210> 129
<211> 172
<212> PRT
<213> Drosophila

<400> 129

```
Met Leu Leu Asn Ile Leu Ala Ile Ile Leu Val Phe Val Ile Ser Ser
 1               5                  10                  15
Gln Ser Glu Ala Val Ile Pro Gly Cys Asp Tyr Phe Asp Thr Val Asp
            20                  25                  30
Ile Ser His Ile Pro Lys Leu Asn Asp Ser Tyr Ala Tyr Glu Glu Leu
        35                  40                  45
Ile Ile Pro Ala His Leu Thr Gly Leu Tyr Thr Phe Arg Gln Leu Ala
    50                  55                  60
Asp Gly Ser Gln Glu Pro Val Lys Ser His Leu Arg Ala Cys Ile Cys
65                  70                  75                  80
Lys Leu Lys Pro Cys Ile Arg Phe Cys Cys Pro Arg Asn Lys Met Met
                85                  90                  95
Pro Asn Ser Arg Cys Ser Asp Gly Leu Thr Glu Asn Leu Lys Arg Ile
            100                 105                 110
Asn Pro Tyr Leu Lys Ile Thr Leu Glu Asp Gly Thr Ile Gly Lys Tyr
        115                 120                 125
Tyr Leu Leu Thr Asp Met Ile Val Leu Arg Tyr Glu Phe Arg Tyr Cys
    130                 135                 140
Glu Lys Val Val Ser Val Gln Glu Asp Gln Tyr Lys Leu Tyr Glu Val
145                 150                 155                 160
Lys Ile Ser Ile Phe Leu Lys Tyr Leu Ser Ile Ile
                165                 170
```

<210> 130
<211> 6552
<212> DNA
<213> Drosophila

<400> 130

```
agtatatacg agtatatcct gtacgactat atatcaaacc gaaagtcgaa ctgaaagtgt 60
cctctaagcg ggactttcat cgaggtcggt gcccctcatt atcgtattca tttgtggggc 120
tcatttgatt cattgtgtct gttcgtgtgt gttggcgtgt tactctgtta agttggatgc 180
aattctgtcg ccgatggttg tggttttgtg tccatggccc atcaatgggg atgctccatc 240
catgtggatg ggcggtcagc caactaactt ggaacattcg tatgtgatgt gtctgtggtg 300
gagggcgccg acctcgaggc ccgctagcta ctgtatccaa ttgaactgtg atttagctaa 360
ctaactctag gtaccgacac ctcgtgtcgg cgtggaacaa tccggcctct aaggatatga 420
atttctttac actcgcttac ctaatcgtac tgtgggaatc ggaaggaagg ttctgtgtgc 480
```

cgtgggtttt ggttttgggt gggaggaaaa gcttgctgct cgtgatatgt gataggttgt 540
gggagcagcc cactttgaag gacactactc tattctggta aagcgaatcg tatccttggc 600
gttttgctta cgtctacttt ggaatcataa tcgcgtgctt gctgtgtgtt gattatatcg 660
ctacttgggg tttgcttaat gtaattgtta aaaattgtac ctagcctaca gcttgcaatc 720
caccgttcat gctccacata gcatatatgt atgtatgtag tttagtggta tgtacagatc 780
tatttacata agtagctagt agaaccttaa agctattggg aaacgcaacg aacgacaaac 840
gaaacgaaac tggagcccaa ttgagatctc tgctgcaacg atggcctcca tttggcacac 900
tttgccatcc tctcccgggc tttaatccat tccacaaggg acccatcatc caattgccta 960
ttgctataac tattgctggt gacgtgctcg ccgttgataa tcaccccatg tgcatgatct 1020
tcttgaaggc cttgcgaaac tcgggattga atattgtgta gataaccggg ttcacaaagc 1080
tgttgatgta gcccagccag gtggtcatca tgtaggcggt gagcccaggt cggcagtctt 1140
tcttgaactt ggcgcacatg gcgtccatga tgttgcagct gaagaacggc agccagcaga 1200
acaggaagac acctgtttga gggtttgaag gttatatata tgatggcata tggattagaa 1260
tatggcagag cgcaaggacg tggtccaaaa cggctggggg aggtggctca acttgaactt 1320
acccaaaaca atggccaatg ttttggtggc cttgcgctcc tttttggccg acgatttttc 1380
gcgtttcttt ttcgaggcct tgtgcacctt gtatatcgtg aatctgtgcg tcacgcgaag 1440
gataataggc catggtgtgg tatatattgc attttccaca ggcaaaaaag gcaaatgatg 1500
aagagaggta gagaaggcga gaaggaaatg gggtggcatt agcatcgact ggccaacaag 1560
agcccaaata actaataact gtaattaggg gcaatggacg gcgttactca ccttgacgcc 1620
tgcgagttct tcttatcctt gcgcgaggac gttttcgatg tagtcgatag cgtcgagtcg 1680
ttgcgtgcca aagtcatggc ctcctgcact ccgtaccgca caaaggacaa tggcttcatg 1740
gcgacgctgc gaacggaagt cggtcggatt cggaatgaga acttaactgc ggcattacgt 1800
actactagtc actattttg ggcttgccgg attgcctaca ttgccaacat gcggcttcta 1860
ttttggggt ttcgaggata aagtggcagg tttctcgcta aacactacgt aataggtctg 1920
tgatgaaaac gtgagctctt ttttgctaag caacttaaat caagcgattg tgcaagaatg 1980
ccaccaaata atcaattgat gatttcataa gcctgctgat atgtaaacat gccgacaata 2040
attatgaaaa gtagtggcaa gtggcagaac taatactgca aatcactatt taaatgtgca 2100
attggatttc aatctgtttg aaaaatgtaa tgtgttaaaa aaaaaaaaaa catgtaatag 2160
ttgacaacaa taacccgaca tatttaatta aactttgcca tttgtttacc aattaatatt 2220
tggaagcaaa gatgttgaca tttttgacgc ggtcaaactc tttgtctgtt tttattgaac 2280
gaaaaaagcc atgcgtgtta tacgcattta ttgttttgtt tataaataat gacaattgat 2340
gtgtgcaatt taagcaattt taatatgcta cgtaatctag tgtaagtaga taaaaaaaaa 2400
cattgcgcaa aaagagcact gcctatagct gaaaaaagta ttatattatt caaaaactat 2460
gtacaaaaac ttgaatggga atgctttatt ttgtcgggct gcaaaaacag ttcattaaac 2520
ttttaaactg ttgcgaaaag tttcatttta cctatggtac aatcaaaaag aatgttgttc 2580
aatgatttag taccaatgtg gtacagatat tggacggtat tttgtttttt tttttttct 2640
aaatattagc aatttctcga atttccgata taattgccgg aactttgatt ttaaaacaat 2700
aacttggatg aggcattatg tttttatccg tgcacgaact tctactacgc cccaactac 2760
tggggatgct atttctaagg tggcaggctt acggatctga atgccaaata cctggagtta 2820
tcgacaccga ctgacctgag ggcgggttcg ccgcgcttcg gcgagtcgtc gccggtgggg 2880
cgccgctgac tgctgacgct ggatcgctgg agggtggcgc cactgggcag cggactgtcc 2940
ggcggcgagc cgctgctccg cggagcggct gatgtggcag cagaggcgga ggcgggcgcc 3000
acgcctgcaa ggacatcgtc aacgtttgag gctgcataac cagaatcatg attaccattc 3060
ggatcggcga cgaccagctg cggctgtgtg gtgatttggg ccacaacgct gctgcgagtg 3120
cgttatggca atgtggagcg atgtgacggt ggggttatcc aattaatggt gccgcgacca 3180
tatccatcaa tcaatcaatc aatcggcgat agacgacgag tgggcgaaga agcagcagaa 3240
tccggcgttg atgaaggacg aaaagatgga aaatcggggt gcaacgtcaa ttaaaggcaa 3300
ggcaagattg agttttaata aagtgacgca ttaattgaat tattgagaat tatttttctc 3360
tgccattggc tgtgatttct tttggcgggt aaagtttcgc ggtcttgttt gccatccatt 3420
tcagcccatc ggttcggttc ggttcggttc ggatcggatc ccatcccttt gtgcatccag 3480
cctcctgtca ttaatggttt catttcatgt ccttattgga ttaagcaggc catcaggatg 3540
ctaattggct gcgcttaacg ctcgcactta atggcggttt cggccaggaa ttgggagaaa 3600
gattcgaatg ccgtggggcg gggacatccg attgcaagga tgatgggcaa gggatccttt 3660
cggcaggacc tgattgaaag tgcagcacgc ttttatgtcg gcaattgagt gccgctgtgt 3720
ccggacgaag cctgccatta gattgacagt tttccatggg catgcaccgt ctgggggtcc 3780
attgctttgc tggtctgctg gtctggaagg aagtgcagtc gagaccactc agcagaaaca 3840
ggcttatcta ttacgtgttt ggggcacgtg ccagggctgt cagggttgtc aggttgcaag 3900
ttgcaagttg caagcggaca aggatatgga cacggaaacg aaactgctga gctccgaact 3960
tttgcagggc catcggagat agcccccagc caaacattcg ttctccaggc gatggagttt 4020
tctgcgaaat tcgccgaaag tttccgcttt tcataccccg ttcacgtaag atgtaagtgg 4080

gtttgcaaat gtacatatat gtccaatccg actccattgg caacgtgttt atattcgtga 4140
agtaacatta gccgccaaaa gaaggtctta cacagccaaa aaaaaacttt gtgaaacgta 4200
tatttccttt ttgtatattc caaatctgat ttaaagctat catgttctac gcatttctac 4260
acagcacatc aaagccccta aatggtatcg taaaatgagt actcacttgc ccgtctcctc 4320
gacaacggtg gccagcatga actcggtgga cttgtcgttc acgatgacgt ggcagtcgtc 4380
gatgtccggc gagatggcgt tctcgtcctc ctcctcgttg gagccgctgg ccgtgttggt 4440
ggcggcctcg tcgggcagaa tcctactggc gtggcgactg ctgtccaggg ccgtctccgc 4500
caggcgccgg gtctgggcaa tgttctcgat gacgctgccg cccgtgagtt ccgataggtg 4560
gggcttgcgg gccgcccgct gcttccgcgc tcgactcctc agcgcctgtg ggtgggtggg 4620
tacgagatgt aaaatacaga aaacagtata ccaaacatgg gacacaggat acagactacg 4680
gacgacggac cacggactac gggctgtggg acacgggatg caggatgcag gatgcaggat 4740
gcagtagaga gcggattagg cactgccatg ctgcgcttgc atcctacgga tattcgcccc 4800
cagtcagcag caaatgtgca caaacagaaa cagaaacagg aagagaaaca gaagcatccc 4860
caagccagca ttgcagatta tggatctgcg ccagcaatat tcattccact cgactcggag 4920
aaatccgatt ttaattgcgg tcttttaata atagtgcttg aaattttttt gcgcaatgga 4980
aagagcgaca gaaagaccaa agcgcggaaa attcgctcta ggaattaatg tttgtgggaa 5040
gcttggagtg ttcaaaggaa ttaacatgtt tggaatatat cgtatatata catatatcgc 5100
tcgctactca ccttgaaaat gttccagtag agaaacacca tgatgatgca gggtatgtag 5160
aagctgctca gcgaggagta gagtataaag tcggcgttgt agaaggcgca tacatcgggt 5220
tcgcggttgg gcgtgttgtt gaggcccaga actatcggcg agccgatggc agccgatatg 5280
gcccacacca gcagtatcgt aaggcaaacg cggcggctat ttttgtgctt ggcgtacttt 5340
attggctgtg tcacagcgat gtatctgtgg gatatatccg tggaatgtgt gcttgggtca 5400
tatgggtctc cgaaataggg gccgctgtga actcacctgt ctatggagat ggcaactaag 5460
ttgaatatcg atgaagtaga gcatatcaca tccatggcta tatagaaatc acaaacgacg 5520
tcaggcaggg cccaggctcc atttacctgc aaaacaaaaa gaaggagggg gggggggcgt 5580
caggtaagaa cgaggatgtg tatgtcgatg cggatgagga gcacacttca tttgttcggc 5640
acttaatcga ctcgggaatg gagctctatt ggaatgccca tcctttgtcg aatattaata 5700
acaatatttt cccatttccc gtgctgctgc cacaaattag ttttagactc tatctcttca 5760
tcgacgtctc cggcagcatc tatgttatgg cccaacccga atgcggactc cataaaaacc 5820
ccaacaaaaa cgtcaggaat ttcccaagcc aaacaatttg tagcacacaa agcacataaa 5880
aaaaggccga gtgggccacc agttcggttg gaacttaatg tggatatgtt tgggtgaaca 5940
cccaacctgt cgtggttgca ctccccatag cccgtgtacg aggtatacat aatgtagcgc 6000
caaccgattc ggtattagaa tggcatcgat agctttttgg atgtgaatac tattagtaag 6060
tagaggtagc tgtcacatag gtatttgttg actgatctgc gggtaagagg tatcgcatag 6120
tcgaatacac tcggcggagc cagtttccca gaatcagtag tttaacattt taagaagtca 6180
atatttgaat aatactacta ctacatatat tgtacttata tgccaccaat gtgtgtgtgc 6240
taactaaaat gcatgtagac ggtatcgaat agtcgaagtt aatcgaccct aacgtactat 6300
cttgtttttc ttcttcaatc ggttttgcgc ctaaattata caaacacatt gaaacttgaa 6360
gcggggctca acgcgtggga agagtgtaaa atattgatac acacgcaacc acaccacacg 6420
aaagtaaaag tgaaagtgaa agtcaagtga aagcagcata gaaatgtgct aaaatagctg 6480
ggcagtcgca gacttggaat ttcttagaac tcggatgaag actccttagc cggagctaag 6540
cgagcataac aa 6552

<210> 131
<211> 1392
<212> DNA
<213> Drosophila

<400> 131
ttgcaggtaa atggagcctg ggccctgcct gacgtcgttt gtgatttcta tatagccatg 60
gatgtgatat gctctacttc atcgatattc aacttagttg ccatctccat agacagatac 120
atcgctgtga cacagccaat aaagtacgcc aagcacaaaa atagccgccg cgtttgcctt 180
acgatactgc tggtgtgggc catatcggct gccatcggct cgccgatagt tctgggcctc 240
aacaacacgc ccaaccgcga acccgatgta tgcgccttct acaacgccga ctttatactc 300
tactcctcgc tgagcagctt ctacataccc tgcatcatca tggtgtttct ctactggaac 360
attttcaagg cgctgaggag tcgagcgcgg aagcagcggg cggcccgcaa gccccaccta 420
tcggaactca cgggcggcag cgtcatcgag aacattgccc agacccggcg cctggcggag 480
acggccctgg acagcagtcg ccacgccagt aggattctgc ccgacgaggc cgccaccaac 540

```
acggccagcg gctccaacga ggaggaggac gagaacgcca tctcgccgga catcgacgac 600
tgccacgtca tcgtgaacga caagtccacc gagttcatgc tggccaccgt tgtcgaggag 660
acgggcaaca gcgttgtggc ccaaatcacc acacagccgc agctggtcgt cgccgatccg 720
aatggtaatc atgattctgg ttatgcagcc tcaaacgttg acgatgtcct tgcaggcgtg 780
gcgcccgcct ccgcctctgc tgccacatca gccgctccgc ggagcagcgg ctcgccgccg 840
gacagtccgc tgcccagtgg cgccaccctc cagcgatcca gcgtcagcag tcagcggcgc 900
cccaccggcg acgactcgcc gaagcgcggc gaacccgccc tcagcgtcgc catgaagcca 960
ttgtcctttg tgcggtacgg agtgcaggag gccatgactt tggcacgcaa cgactcgacg 1020
ctatcgacta catcgaaaac gtcctcgcgc aaggataaga agaactcgca ggcgtcaaga 1080
ttcacgatat acaaggtgca caaggcctcg aaaaagaaac gcgaaaaatc gtcggccaaa 1140
aaggagcgca aggccaccaa aacattggcc attgttttgg gtgtcttcct gttctgctgg 1200
ctgccgttct tcagctgcaa catcatggac gccatgtgcg ccaagttcaa gaaagactgc 1260
cgacctgggc tcaccgccta catgatgacc acctggctgg gctacatcaa cagctttgtg 1320
aacccggtta tctacacaat attcaatccc gagtttcgca aggccttcaa gaagatcatg 1380
cacatggggt ga                                                      1392

<210> 132
<211> 444
<212> PRT
<213> Drosophila

<400> 132
Met Asp Val Ile Cys Ser Thr Ser Ser Ile Phe Asn Leu Val Ala Ile
 1               5                  10                  15
Ser Ile Asp Arg Tyr Ile Ala Val Thr Gln Pro Ile Lys Tyr Ala Lys
            20                  25                  30
His Lys Asn Ser Arg Arg Val Cys Leu Thr Ile Leu Leu Val Trp Ala
        35                  40                  45
Ile Ser Ala Ala Ile Gly Ser Pro Ile Val Leu Gly Leu Asn Asn Thr
    50                  55                  60
Pro Asn Arg Glu Pro Asp Val Cys Ala Phe Tyr Asn Ala Asp Phe Ile
65                  70                  75                  80
Leu Tyr Ser Ser Leu Ser Ser Phe Tyr Ile Pro Cys Ile Ile Met Val
                85                  90                  95
Phe Leu Tyr Trp Asn Ile Phe Lys Ala Leu Arg Ser Arg Ala Arg Lys
            100                 105                 110
Gln Arg Ala Ala Arg Lys Pro His Leu Ser Glu Leu Thr Gly Gly Ser
        115                 120                 125
Val Ile Glu Asn Ile Ala Gln Thr Arg Arg Leu Ala Glu Thr Ala Leu
    130                 135                 140
Asp Ser Ser Arg His Ala Ser Arg Ile Leu Pro Asp Glu Ala Ala Thr
145                 150                 155                 160
Asn Thr Ala Ser Gly Ser Asn Glu Glu Glu Asp Glu Asn Ala Ile Ser
                165                 170                 175
Pro Asp Ile Asp Asp Cys His Val Ile Val Asn Asp Lys Ser Thr Glu
            180                 185                 190
Phe Met Leu Ala Thr Val Val Glu Glu Thr Gly Asn Ser Val Val Ala
        195                 200                 205
Gln Ile Thr Thr Gln Pro Gln Leu Val Val Ala Asp Pro Asn Gly Asn
    210                 215                 220
His Asp Ser Gly Tyr Ala Ala Ser Asn Val Asp Asp Val Leu Ala Gly
225                 230                 235                 240
Val Ala Pro Ala Ser Ala Ser Ala Ala Thr Ser Ala Ala Pro Arg Ser
                245                 250                 255
Ser Gly Ser Pro Pro Asp Ser Pro Leu Pro Ser Gly Ala Thr Leu Gln
            260                 265                 270
Arg Ser Ser Val Ser Ser Gln Arg Arg Pro Thr Gly Asp Asp Ser Pro
        275                 280                 285
Lys Arg Gly Glu Pro Ala Leu Ser Val Ala Met Lys Pro Leu Ser Phe
    290                 295                 300
```

Val Arg Tyr Gly Val Gln Glu Ala Met Thr Leu Ala Arg Asn Asp Ser
305 310 315 320
Thr Leu Ser Thr Thr Ser Lys Thr Ser Ser Arg Lys Asp Lys Lys Asn
325 330 335
Ser Gln Ala Ser Arg Phe Thr Ile Tyr Lys Val His Lys Ala Ser Lys
340 345 350
Lys Lys Arg Glu Lys Ser Ser Ala Lys Lys Glu Arg Lys Ala Thr Lys
355 360 365
Thr Leu Ala Ile Val Leu Gly Val Phe Leu Phe Cys Trp Leu Pro Phe
370 375 380
Phe Ser Cys Asn Ile Met Asp Ala Met Cys Ala Lys Phe Lys Lys Asp
385 390 395 400
Cys Arg Pro Gly Leu Thr Ala Tyr Met Met Thr Thr Trp Leu Gly Tyr
405 410 415
Ile Asn Ser Phe Val Asn Pro Val Ile Tyr Thr Ile Phe Asn Pro Glu
420 425 430
Phe Arg Lys Ala Phe Lys Lys Ile Met His Met Gly
435 440

<210> 133
<211> 2970
<212> DNA
<213> Drosophila

<400> 133
tccatgccag gcctggatgc acatgggccg gatggggatc aggtatatat cgagagcgtc 60
gagagttcgg tttgcgtttt gcccaccaag acgaaaccca aaaaggtcgc cttttacggc 120
agcaatggtc agcgctacac tttcctgttc aagggcatgg aggatctcca tctggacgaa 180
cgcatcatgc agttcctgtc catctcgaat gccattatgg cctgccgcag cgacgcaccc 240
ggcaatggct gctatcgtgc ccaccactat tcggttatac cattgggacc gcaatccgga 300
ctgattagct gggtggacgg ggtcacacca gtctttgcgc tctacaagaa gtggcaacag 360
agacggtccc aagttgctgg gaatgcggga gcaggtgctg tggcaaatgt cccgcgacgt 420
ttcacagact tgttctacaa caaactctcg cccctgttgg ccaagcacaa tatgcaggtg 480
agcgatccac gacgccaatg gccaatatcc gtcctgctcc aggtgctcga cgaattgtca 540
caggaaacac ccaatgatct gcttgcccgc gaactgtggt gccaggcggg aaatgccgcc 600
gagtggcggc aatcggtgcg tcgctttgtc cgctgcatgt cggtcatgtc gatgattggc 660
tatgtcattg gcctgggcga tcgtcatttg gacaacgtgc ttatcaacct gggaagcggc 720
gacatcgtgc atatcgacta caacgtgtgc ttcgagaagg gccgcaccct gcgcatcccc 780
gagaaggtgc ccttccgtct tacccagaat ctagtgcagg ccatgggcat cacaggaatc 840
gaggtaagac tattgctcat cctataccaa tcaattatct caacaaatcg ctgctggggg 900
tgtttaagaa acggtgctcc taaagataat gtccactttt attgctatat ctccatttaa 960
ttctaccaag ttaaagcgtt gggtttccca attccgttga tcactgagtg gccttcggtt 1020
tgatctcctg agatgaggat gtggcctcca tgcggtgctt gctgtccgcc ggcaattgat 1080
cctcgcactg cctcgccgtt cggcagcata atgtttggac aaaggctcgc cgaaagccag 1140
agttcttcca cgagtagata atcggattga gagcagagtt ggcaatggcc agagagaacg 1200
tcgtcttgta gatcatcgag ctgctctggc agatgctaaa tagctgggca atggccacac 1260
agaagtacgg tagccaacag agcgtgaagc agcccatgat gaagaccaca atctggacgc 1320
tcttccagtc tggatggagc agcagatttc gcatggtggt ggcctcgtcc gtgttgtaga 1380
ccaccgactg tcgcaatcgc agcgcctgct tggatgcctc gcgcatgatg cgccagtaga 1440
cgaggaacat acagatccag atgatcacaa aacccggcgt tatcactccg gcaatgtaac 1500
caggcgcgag aacttcgtcg aattcgcacg cctgtgcatc cggccatcga ttccagaaca 1560
ctggcagcag agccaccaat gcacccaagc accaattgaa tatgatgatg ctatatgcca 1620
ctcgacgggt catgtatcta cggtttcaaa gggtttcaca tatgtagcca ttgaattagg 1680
atttatcgcg ttcgcttaat accttacctt ctatagtgca gagcatagac gactgctata 1740
taccgatcca ccgcaatcga gatcagagtc aacatggaca cgcagcaggc gcagatgagc 1800
aggaagaagc gcagcaggca gggacctctc atggcgccaa tgtccgagcc catgtaaaat 1860
accagatgat agggcagtgc cagtcccaca cagaagtcgg agacggccag cgacaggatg 1920
aacagattgg agatgaccga tcgcaggtgg cggcaggtgc gcaccgccac aatggtcaga 1980
atattgccgc ccaggataag gacaaacaag aaggcattga tggctagcca cagcagatgg 2040

```
ccggcaccga aactggtcgc ctggagggtt atttccgttg gcgtttccgt ggcgttctgg 2100
tgctgccaca ggaacgggac gacacccaac gcttccgggg agcttcgata atcccaagcg 2160
ggcgatgagc tgctggtaaa tgccatgagc cacgatggta aacgaagcgc acgcgacgcc 2220
agtggttggc tatctggcaa ccgtgggagc tggagttgtg ctgacttttg gccaggactt 2280
cggctcggct cggatcggat tcagattcaa ttgggtttcg gctcggaccg tgggcgtcgt 2340
ttgtttacgt ttcgctggtg gctttacgac atcgtcatcg cgccagttaa ctgatagcgc 2400
gcccagatcc accgtgagaa ctgcaatccg ttgataagcc ctccactttg gcccaaccac 2460
cattcaccac cctcatacct aaccactcac ccacccatcg agtcaatcca accactccaa 2520
ccagccaaac tatccattgc cccgagtat ttcgagtgtc cggactcccg ggcccaatgc 2580
atggccaaac agatcgctgg gggttttagc agattagccc agctccgcgg accaaaatag 2640
attgataaaa ggccagggcc agcaactggg atacgtggta aacagataca tcattgatat 2700
tcccaaattt cctcgataca ttgcaagcga agaaactata tgtaaatagt gatatcatca 2760
taattgttat aaatccgaga ccaacagtac aaccagtact attaggattg acatggttat 2820
ccactcttgg ttcttttaga attagaaatc aggagaatat atagtatata taggtacaac 2880
tctctgatct aatcccatat ttgaattatc tttaagatga gatattaaag atttcatgga 2940
tcagaacaga ccaggagttt gctaaggatg                                  2970
```

<210> 134
<211> 1116
<212> DNA
<213> Drosophila

<400> 134

```
atggcattta ccagcagctc atcgcccgct tgggattatc gaagctcccc ggaagcgttg 60
ggtgtcgtcc cgttcctgtg gcagcaccag aacgccacgg aaacgccaac ggaaataacc 120
ctccaggcga ccagtttcgg tgccggccat ctgctgtggc tagccatcaa tgccttcttg 180
tttgtcctta tcctgggcgg caatattctg accattgtgg cggtgcgcac ctgccgccac 240
ctgcgatcgg tcatctccaa tctgttcatc ctgtcgctgg ccgtctccga cttctgtgtg 300
ggactggcac tgccctatca tctggtattt tacatgggct cggacattgg cgccatgaga 360
ggtccctgcc tgctgcgctt cttcctgctc atctgcgcct gctgcgtgtc catgttgact 420
ctgatctcga ttgcggtgga tcggtatata gcagtcgtct atgctctgca ctatagaaga 480
tacatgaccc gtcgagtggc atatagcatc atcatattca attggtgctt gggtgcattg 540
gtggctctgc tgccagtgtt ctggaatcga tggccggatg cacaggcgtg cgaattcgac 600
gaagttctcg cgcctggtta cattgccgga gtgataacgc cgggttttgt gatcatctgg 660
atctgtatgt tcctcgtcta ctggcgcatc atgcgcgagg catccaagca ggcgctgcga 720
ttgcgacagt cggtggtcta caacacggac gaggccacca ccatgcgaaa tctgctgctc 780
catccagact ggaagagcgt ccagattgtg gtcttcatca tgggctgctt cacgctctgt 840
tggctaccgt acttctgtgt ggccattgcc cagctattta gcatctgcca gagcagctcg 900
atgatctaca agacgacgtt ctctctggcc attgccaact ctgctctcaa tccgattatc 960
tactcgtgga agaactctgg ctttcggcga gcctttgtcc aaacattatg ctgccgaacg 1020
gcgaggcagt gcgaggatca attgccggcg gacagcaagc accgcatgga ggccacatcc 1080
tcatctcagg agatcaaacc gaaggccact cagtga                            1116
```

<210> 135
<211> 371
<212> PRT
<213> Drosophila

<400> 135

```
Met Ala Phe Thr Ser Ser Ser Ser Pro Ala Trp Asp Tyr Arg Ser Ser
 1               5                  10                  15
Pro Glu Ala Leu Gly Val Val Pro Phe Leu Trp Gln His Gln Asn Ala
            20                  25                  30
Thr Glu Thr Pro Thr Glu Ile Thr Leu Gln Ala Thr Ser Phe Gly Ala
        35                  40                  45
Gly His Leu Leu Trp Leu Ala Ile Asn Ala Phe Leu Phe Val Leu Ile
    50                  55                  60
Leu Gly Gly Asn Ile Leu Thr Ile Val Ala Val Arg Thr Cys Arg His
65                  70                  75                  80
Leu Arg Ser Val Ile Ser Asn Leu Phe Ile Leu Ser Leu Ala Val Ser
```

85 90 95
Asp Phe Cys Val Gly Leu Ala Leu Pro Tyr His Leu Val Phe Tyr Met
100 105 110
Gly Ser Asp Ile Gly Ala Met Arg Gly Pro Cys Leu Leu Arg Phe Phe
115 120 125
Leu Leu Ile Cys Ala Cys Cys Val Ser Met Leu Thr Leu Ile Ser Ile
130 135 140
Ala Val Asp Arg Tyr Ile Ala Val Val Tyr Ala Leu His Tyr Arg Arg
145 150 155 160
Tyr Met Thr Arg Arg Val Ala Tyr Ser Ile Ile Ile Phe Asn Trp Cys
165 170 175
Leu Gly Ala Leu Val Ala Leu Leu Pro Val Phe Trp Asn Arg Trp Pro
180 185 190
Asp Ala Gln Ala Cys Glu Phe Asp Glu Val Leu Ala Pro Gly Tyr Ile
195 200 205
Ala Gly Val Ile Thr Pro Gly Phe Val Ile Ile Trp Ile Cys Met Phe
210 215 220
Leu Val Tyr Trp Arg Ile Met Arg Glu Ala Ser Lys Gln Ala Leu Arg
225 230 235 240
Leu Arg Gln Ser Val Val Tyr Asn Thr Asp Glu Ala Thr Thr Met Arg
245 250 255
Asn Leu Leu Leu His Pro Asp Trp Lys Ser Val Gln Ile Val Val Phe
260 265 270
Ile Met Gly Cys Phe Thr Leu Cys Trp Leu Pro Tyr Phe Cys Val Ala
275 280 285
Ile Ala Gln Leu Phe Ser Ile Cys Gln Ser Ser Ser Met Ile Tyr Lys
290 295 300
Thr Thr Phe Ser Leu Ala Ile Ala Asn Ser Ala Leu Asn Pro Ile Ile
305 310 315 320
Tyr Ser Trp Lys Asn Ser Gly Phe Arg Arg Ala Phe Val Gln Thr Leu
325 330 335
Cys Cys Arg Thr Ala Arg Gln Cys Glu Asp Gln Leu Pro Ala Asp Ser
340 345 350
Lys His Arg Met Glu Ala Thr Ser Ser Ser Gln Glu Ile Lys Pro Lys
355 360 365
Ala Thr Gln
370

<210> 136
<211> 7659
<212> DNA
<213> Drosophila

<400> 136
gcctaaacga taagattaga aaatggaaat ggaaatgaag acagccacct gacgcaatct 60
cacatgccag aaatttgtcg tttgccatcc aaggccccga tttccgaaat cgaaaatcct 120
tgatatccaa ctggacatga ccagtttccg gaggaatgca ctccagttac acaactggca 180
caaacatttg tttgctgttt tctgaccaaa aacaagccac acttatgctg gcactttttt 240
gttgacagaa atatgcaagc agacacaagt aaatatttat ttgattattt attttaaaga 300
aatgtgcaag ctcaccgaat aaatattata ttttaaaatc aataagtaga aacctaaaca 360
aatacaagct ggtgaagcga ttgctctgca aattccacat atttaaacta ttaatttaca 420
aaaataaatg tggcggtcaa tttccagctg aacagttctc aacaattaat gaaaataatt 480
gccactgata ataatagttt tttttattat tcgattacat actttaaata ttttttttagt 540
ttcccaatta gaaaactttt ttttaacaaa atatttaata ttccgtaaca gtatttatgc 600
ttaattacca gatgacagtt tttatttctt ttgactatac ttacattatt ttgaaaaaat 660
agcaattatt tttagacttc tgccacgaat gcctttaatt aatagctgat gcaatcgctt 720
aaactcaatt aagttacaca cgcaacccac atcttctgtg taaaaaaaaa tcaatgtatt 780
cgaccataaa taattccgct gcaatcagtg caacagtatt cagtcctgga actttaatag 840
cgccaagctg ttggaaaacc aattaaattc cctttcaaat ggagcgtaaa acgcaaaata 900

```
cttacgtatg atgggggcgc cacacactcg cacacgcatt tatgagctcg gcaaaaccca 960
tttgattgac taacgaaact tatttgccac tttatttcag cggacgcagt gcagcagttc 1020
cacgacgacc atgagcgacc agattggcaa ccccaatgca acattcagtg gcagtggcag 1080
tggcagtggc accaacgtcg cctccatcgc cgaaagtgtg gcggagagcg gaccggactt 1140
cgatgcactt cgggccgcct gcgaaacgcg gctaaatgcc agcggtcaac tggcgggtga 1200
gtcgtgcact gcgagaaaat atatcacaca gcacataact attaaagttt atgcagttat 1260
tttatttgat gcagttgtaa aacgctcacc agtaaatttc atttaaattt tagatgtttt 1320
tataaacata aatcaaaata aatttttttg cactgtttta cttacattcc ctttctttat 1380
gccagttatt tgtctttact taagcgaaag aaaaaattca ctctgaaatt gtagcactag 1440
ggaaatataa cagtttttgg tatttcacta catatttttg cccgtgcttc tggctctggc 1500
caactgcaat ttcggcgtcc attatgcacg agtaacttta aattgctgac ctaaccactc 1560
ccctttcctt ttggcccctt ggcaaacagg ctccggagga ccaggcgcag aagcaggaac 1620
ccactgtgcc ggcacctttg atggctggct ttgttggccg gatacggctg tcggcacttc 1680
cgcctacgaa ctctgcccgg acttcatcac gggattcgat ccagcaagta agtgcgccgg 1740
ggcaggagct gaattgttta ttaaattcca caaagctaca tccgtttcgg ggtcagctgg 1800
ggatctgcat atgacacgaa aaaatatata tatatacata catatatcac atgtatgtat 1860
gttacatata ggttaaatag catgccaaag ggaaatgtgt ttgcacgagg caatgcggtt 1920
ggataagttg gttgggttgt ttggtaggtg gagtgtgcca gtctaccctc taattgggct 1980
catccccgag ctcatcgtta tatacacata aatggaacag aaaatacgct ccaagccagg 2040
ccaaatcaat tggaaacttg ttatgtttat gctttgattt cgtgtcgtat ggaaacatca 2100
agtattcaca aaaccacacg ggctaaaggc aaagctcctt ggagctccct gggccactaa 2160
attatccaaa agagctcgac aaaatgctaa aatgtttcgc ctgatttggc acaatttgcc 2220
agagtctaac cgaaatcaaa tgtaaggcag aaaacacggc caaattgatt gaggctgctg 2280
ggcagagtag aagtcaaata tgtgtgtggg tatttaaaga tgctggagca cgaagaaagg 2340
ggatttaaat tttatcatgt ttgtgaattt catatttttt taatcattaa tcagtgttct 2400
tggcaaggtg ttgatgacgg tcataataac ataactgaag aagaaaaccc tttgaatagt 2460
ttcaccaact cattgtcgta tcattaacat atcataaata aaaatatcaa aggtgacttg 2520
aattatttct caactttctg ccacactgcg tgtatttcat tcaattataa atatatatat 2580
atataaatat aaatataaat taatcaactg attggccaaa cgcacaaaga aaactaatta 2640
attatgcttg gcattaaagc aaacgtttaa aaacaaagcg cacgctctag aaaatcgtca 2700
taaatgtgga taaatatttg aaccaattta tggcataaat catgcgctgc atgataattt 2760
gcattaatgg tatacaatat gtagccagcc aacttaaacc tcaacctcgt tttggccatg 2820
accagggcaa cgaaaatagc gagagttgaa tgactttttt tgttatgaca gagcggagaa 2880
aaacaaggcg agcgtcaaga aagtgctact tttgctacat ttagcttgtg gttagcatgg 2940
cgccaccccc tttttcatcc tttttgcatt tttttttttg ccttacgggg tcgccgccag 3000
cttgtacaag attatgtgct ttggcactgc gccgccatta acttttatgg cacttccggt 3060
cgccggatgt cgatgtggtg agctgtggtc gagaactccc tcaccccccg gctacttatc 3120
tgccacacag cgcgtatacg tgattttctc ggtctgcgtg ccaaacggcg ataaatggat 3180
tcttggccaa agatcatcgg tcgaaaaaaa aaagaaaacg gaagagccgg aaaaacgaac 3240
gacctgataa aactcacctt cgaagttcct aacaaaaatt ctccagtaat ttttctacaa 3300
gtgccactcg gtaaattgca ttgcaaactt ggtaatctac aatcgaatcg aatggaaagt 3360
ttttgtcacg aacttcgaat tgattcgact taaacacaca catgacccaa acaacagaat 3420
gattatgcaa ttttcaaatt gatatgcttg caattcgaat gcgaaccgca attcttttgc 3480
ggttaacttg gtggcagatg caataaataa tggttaattc aattgaatgc tcgggccgac 3540
cgcaaaatgt aaatatgagc cggaatattt atgcagtagt caatcatgaa atcacacata 3600
tttagcaagt gtaaatattt gtgtggtgca caggatgtga aaatgtttgc tagaaaaaac 3660
ggtaaagaat taatcgtata tacaatattt attgaggagc tgtaataagc agagtaaaca 3720
gcatggttaa ttatttttta tttatcgcaa actgacattg ctttaattat gcaggatatg 3780
cccacaagga atgcggtctc gatggcgagt ggttcaagca tccgctgacc aataaaacat 3840
ggtccaacta cacaacctgc gtaaatctcg aagacctcaa cgtaagccga tcccatacat 3900
aaatatccta tatatccttt atgtaattaa ttaatcgcaa cactatatat ccagtggagg 3960
cacacagtga atctgatctc cgaggttggc tacggcacct cactgctggc cattctgctg 4020
tcgttggcca tattgggtta tttcaagtga gtatggccaa gggctttgca ctccttcatc 4080
ttggccacaa gttgacagct tcgttgggag atggcaaaaa cacttagggg taatgaggca 4140
cgctggtgac ttatattgcg acttatatta tgacgactgg cgccataagt gaaatttact 4200
gatggcaaac aaatgcgaca gtggcttaat caactgccaa cctggatgct ggatgcccaa 4260
atgctcattg cactcattgc tcatttaatg cggcacaatt gacgccggaa atgggaaaat 4320
acattgtgtc cccatctcct ggtgactcga agcaggatat cgattaattt cccagtgcct 4380
gcagaccaat tagtgctctg ggtttgttta tcgcacattc gaatggaact ggggctcaag 4440
gataaaggat aatggatgtg cggactgagg ggcgagatcc catagacggt tcggccaact 4500
```

```
gacagaactg ttttaaattg tgtgccaaag gcaaataaaa tgggtccaga tacgtgtggg 4560
aatttggtct gccctaaaaa agtgaataat taaaagttta atgacagagc aatgggcaac 4620
accagccccg gaaaaacgca ttcctcaaac agctgtaatg aaaataccag agttttcatt 4680
caaataaaat ataagagtat aaaggaaatt gagtcgttgt gagcggccaa aaataaataa 4740
agaagtatta ttataaaata taccatcagt atattctcag tcaataattg tattgcttaa 4800
agaaagttta gaagtgaaaa gtagtatata aaataattta tttagacaga agtaaattaa 4860
gtaaagtaga gaagggatta acgaaaacaa tctaacctat tgaattaaat tgctgaaaag 4920
aaatcaaact agattcagtt tggccagatg tcagcaatca atgcaattga gcaaaggaga 4980
aacggaaaat gggaaaatgg acaaatgaac tttgtccttt tggtcgcagg actttttctg 5040
cctctttcgg ccctatccaa gctgcctcac atgcccctgg ccccgaatgc aaatgaagcc 5100
atagacaaac ttgcaaattt ggccataaaa acttgccaca aattgccggc caaagtaatt 5160
cgattcgcat atttgtcttt cgattgaaca gatccctgaa gtgcgcccgc atcacgctgc 5220
acatgaatct gttcgcctcg ttcgctgcca ataactcgtt gtggctggtc tggtacctgc 5280
tggtcatgcc gaattcggag ctactgcatc agagtccggt aagtgtgcac acctggcgaa 5340
tgcgtaatgg tgaatcaacc tcggcttccg atttccatat tccatactcc gtgcgaacac 5400
ccagatgcgc tgcgttgccc tgcacataac gctacactac ttcctcctgt ccaattactc 5460
ctggatgctc tgcgagggat tctatctgca caccgtcctc gtcgccgcat tcatttccga 5520
gaagaggctg gtcaaatggc tcatcgcatt cggctggggc tccccggcca tcgtcatatt 5580
cgtctatagc atggctcgcg gtctgggcgg cacgcccgag gacaatcgtc agtaagtgaa 5640
taatgaatgg gtacagctta caccgaaaga aaatgaggaa tgccaatagt tatcaaaagt 5700
gaccgacaac agacaaggtg taaccatatt gtgttgtcta gaagtatcct agtttaatga 5760
aatcccacag acttgagtat cacagattga tatcaagata tttctgtgta gctaagatat 5820
tcagtctgtc gagctcttga ccaaatgcta atcaaatcaa tcagcaattc acccactgac 5880
tcatctcgca ttccttcgcc aacttgcagc tgctggatga accaaaccaa ctaccaaaac 5940
attcttatgg tgcctgtgtg tatctccatg ttcctgaacc tcctgttcct gtgcaacatt 6000
gtgcgagtgg tcctttgaa actgaatgcc ccggccagta ttcagggcag ctgcggtcca 6060
tcgcgaacgg ttttgcaagc atttcggtaa gttttcccgt ccaaaacaaa agaaaaaaaa 6120
aagaactgtt gggggggaga ttgaaacact tataagcgaa gcagagaact ctcaactggc 6180
ccacggggcg aatgcgtaat tatattttgt ttcccttgta tttagactct ttcccatcgg 6240
ccatcaggtc ttttattgtt tattcgctct gcactctgga agttgggctc attttacatt 6300
ttgaattacg catacgccgc gttgacttgc atgttaatac acagctgaaa aacgaaaaca 6360
aggcattata attcaccctc tcatcctgtt ttagggcaac gctgcttttg gttcctctgc 6420
tcggcctcca gtacatcctt acgccctttc gtcctgcacc caaacatccc tgggagaata 6480
catacgaaat catctcggcg tttacggcct catttcaagt aagtaaaatg ttttcgtttc 6540
gtttctgttc tcgtttcgtt ttttttcgt ttttttttt tgctttttc tttccggctt 6600
ttgttttgca cgctctcata aaatcacttt acgactccga gctggctaaa aagttatgat 6660
gttagttcgc ctcatggtgc catccaccac tgggaatgca aataagccag cattactgga 6720
ataccgaata ccaaatacca ggactagcag ctcattagat cggggctctt ttcaaatcaa 6780
taatgaaaat acacgccatg tgaaataccc accattggca tttcacttga tctcagttta 6840
ttaaattgct ggccgagatt ggagcgaagt tcccccgttg gaaaaatgta gcaagttgag 6900
catgaaatgc actctcgcac acaggaactc gtaaaagcaa ctcgagcttg tccctcggcc 6960
aactcagtct cgatagacca gcatttacac atggttaaca cacatacata tatttcttgg 7020
ccgcaactta aagacaacga cggcaaactt catccagcta cgttgccttc ttgtttgttg 7080
aatacccaaa gaatctttca gttcatgttt gcagtacaca gataaaataa tttgcatata 7140
ttttcgtgtt tttgtgcacg aaatggtttt tcattacaga ctgttgatat aacagttgcc 7200
atggaggtag aggtgcctaa aagtatgttg cagcttagat ggctttcctt tacaaaattt 7260
caaaagtaat ttctaaggca ccccctagct gttattcgaa aatcttttag ttcgatctaa 7320
aacccaactc aatcggttca aaaaggtagc ccattgaatt cataaacagc cctaaaacaa 7380
caggaaaatc cattatcttg cgagtagaca taaaatcatc ttgggtgcca ctaagctgca 7440
gttgagctct tgactatttt ctataacttg gcaaactgaa cattgtcacc tcggtctgag 7500
cctttttca tcattgccaa tcgataggtg aaatctttta ttcgactgct aaaagtttaa 7560
agccagtcga cgggtacttt tcctttgcat tttcctcttg tcgtatttat tcacttttt 7620
ttatgtcgct gttttccgtt ggcttatttg ctgcaaaga                       7659
```

<210> 137
<211> 1365
<212> DNA
<213> Drosophila

<400> 137
cggacgcagt gcagcagttc cacgacgacc atgagcgacc agattggcaa ccccaatgca 60
acattcagtg gcagtggcag tggcagtggc accaacgtcg cctccatcgc cgaaagtgtg 120
gcggagagcg gaccggactt cgatgcactt cggccgcct gcgaaacgcg gctaaatgcc 180
agcggtcaac tggcgggctc cggaggacca ggcgcagaag caggaaccca ctgtgccggc 240
acctttgatg gctggctttg ttggccggat acggctgtcg gcacttccgc ctacgaactc 300
tgcccggact tcatcacggg attcgatcca gcaagatatg cccacaagga atgcggtctc 360
gatggcgagt ggttcaagca tccgctgacc aataaaacat ggtccaacta cacaacctgc 420
gtaaatctcg aagacctcaa ctggaggcac acagtgaatc tgatctccga ggttggctac 480
ggcacctcac tgctggccat tctgctgtcg ttggccatat tgggttattt caataattcg 540
attcgcatat ttgtctttcg attgaacaga tccctgaagt gcgcccgcat cacgctgcac 600
atgaatctgt tcgcctcgtt cgctgccaat aactcgttgt ggctggtctg gtacctgctg 660
gtcatgccga attcggagct actgcatcag agtccgatgc gctgcgttgc cctgcacata 720
acgctacact acttcctcct gtccaattac tcctggatgc tctgcgaggg attctatctg 780
cacaccgtcc tcgtcgccgc attcatttcc gagaagaggc tggtcaaatg gctcatcgca 840
ttcggctggg gctccccggc catcgtcata ttcgtctata gcatggctcg cggtctgggc 900
ggcacgcccg aggacaatcg tcactgctgg atgaaccaaa ccaactacca aaacattctt 960
atggtgcctg tgtgtatctc catgttcctg aacctcctgt tcctgtgcaa cattgtgcga 1020
gtggtccttt tgaaactgaa tgccccggcc agtattcagg gcagctgcgg tccatcgcga 1080
acggttttgc aagcatttcg ggcaacgctg cttttggttc ctctgctcgg cctccagtac 1140
atccttacgc cctttcgtcc tgcacccaaa catccctggg agaatacata cgaaatcatc 1200
tcggcgttta cggcctcatt tcaagtaagt aaaatgtttt cgtttcgttt ctgttctcgt 1260
ttcgtttttt tttcgttttt tttttttgct tttttctttc cggcttttgt tttgcacgct 1320
ctcataaaat cactttacga ctccgagctg gctaaaaagt tatga 1365

<210> 138
<211> 444
<212> PRT
<213> Drosophila

<400> 138
Met Ser Asp Gln Ile Gly Asn Pro Asn Ala Thr Phe Ser Gly Ser Gly
1 5 10 15
Ser Gly Ser Gly Thr Asn Val Ala Ser Ile Ala Glu Ser Val Ala Glu
20 25 30
Ser Gly Pro Asp Phe Asp Ala Leu Arg Ala Ala Cys Glu Thr Arg Leu
35 40 45
Asn Ala Ser Gly Gln Leu Ala Gly Ser Gly Gly Pro Gly Ala Glu Ala
50 55 60
Gly Thr His Cys Ala Gly Thr Phe Asp Gly Trp Leu Cys Trp Pro Asp
65 70 75 80
Thr Ala Val Gly Thr Ser Ala Tyr Glu Leu Cys Pro Asp Phe Ile Thr
85 90 95
Gly Phe Asp Pro Ala Arg Tyr Ala His Lys Glu Cys Gly Leu Asp Gly
100 105 110
Glu Trp Phe Lys His Pro Leu Thr Asn Lys Thr Trp Ser Asn Tyr Thr
115 120 125
Thr Cys Val Asn Leu Glu Asp Leu Asn Trp Arg His Thr Val Asn Leu
130 135 140
Ile Ser Glu Val Gly Tyr Gly Thr Ser Leu Leu Ala Ile Leu Leu Ser
145 150 155 160
Leu Ala Ile Leu Gly Tyr Phe Asn Asn Ser Ile Arg Ile Phe Val Phe
165 170 175
Arg Leu Asn Arg Ser Leu Lys Cys Ala Arg Ile Thr Leu His Met Asn
180 185 190
Leu Phe Ala Ser Phe Ala Ala Asn Asn Ser Leu Trp Leu Val Trp Tyr
195 200 205
Leu Leu Val Met Pro Asn Ser Glu Leu Leu His Gln Ser Pro Met Arg
210 215 220

```
Cys Val Ala Leu His Ile Thr Leu His Tyr Phe Leu Leu Ser Asn Tyr
225                 230                 235                 240
Ser Trp Met Leu Cys Glu Gly Phe Tyr Leu His Thr Val Leu Val Ala
                245                 250                 255
Ala Phe Ile Ser Glu Lys Arg Leu Val Lys Trp Leu Ile Ala Phe Gly
            260                 265                 270
Trp Gly Ser Pro Ala Ile Val Ile Phe Val Tyr Ser Met Ala Arg Gly
        275                 280                 285
Leu Gly Gly Thr Pro Glu Asp Asn Arg His Cys Trp Met Asn Gln Thr
    290                 295                 300
Asn Tyr Gln Asn Ile Leu Met Val Pro Val Cys Ile Ser Met Phe Leu
305                 310                 315                 320
Asn Leu Leu Phe Leu Cys Asn Ile Val Arg Val Val Leu Leu Lys Leu
                325                 330                 335
Asn Ala Pro Ala Ser Ile Gln Gly Ser Cys Gly Pro Ser Arg Thr Val
            340                 345                 350
Leu Gln Ala Phe Arg Ala Thr Leu Leu Leu Val Pro Leu Leu Gly Leu
        355                 360                 365
Gln Tyr Ile Leu Thr Pro Phe Arg Pro Ala Pro Lys His Pro Trp Glu
    370                 375                 380
Asn Thr Tyr Glu Ile Ile Ser Ala Phe Thr Ala Ser Phe Gln Val Ser
385                 390                 395                 400
Lys Met Phe Ser Phe Arg Phe Cys Ser Arg Phe Val Phe Phe Ser Phe
                405                 410                 415
Phe Phe Phe Ala Phe Phe Phe Pro Ala Phe Val Leu His Ala Leu Ile
            420                 425                 430
Lys Ser Leu Tyr Asp Ser Glu Leu Ala Lys Lys Leu
        435                 440
```

<210> 139
<211> 4185
<212> DNA
<213> Drosophila

<400> 139

```
attcgttcgg tattttcgca tttggcatgt ttgtggctaa gatatcatca attcaacaca 60
atgatcaaca atattcaacg acaaacactc gacttgggac cctggatgct tcatcggcgg 120
cgctcgggaa actgaatttg acccactaag ggaactaggc aaaattgcaa cgctctgttg 180
gtaagcccac caaatcgaag cgttcagggg cggccaaaag gaaaatgaat cattctaccg 240
atttgaacaa cgatcgaaac ttcacactag cctatcattt cagtatcata gttgtatgtt 300
gtatggctaa gaactactgg aaatgtcttt ttgggcttat ttgcttttttg gattttaaca 360
tatttgctgt gttttatatt cgtttagtgg gccttggttt ttcgtttctt aagcgcaatt 420
cgcaaagcgt gcgaagattg aagattttcc gaattttgaa ttctcatatt tcggcccgtc 480
gtcaagtttg gcaaattgtt taataaacgc tacacgattt cattttgtaa aattctaaat 540
gctacttaat atggactttg aaaaaggttt atacttatgc aatctttaga taataatcct 600
gctgctcgag tagcatatga ttttcgtata ttgtttttgt tgtagcaagt tcgaattgta 660
aaccggaaac ggaaatgttg ctcttgccta gtgaattgtc tcttgtctac gaattagaat 720
tagaattaga tcgctgctac atatgtatat gtatatggct agagttaagt acaagtatat 780
atatacgtgg taagctcata aatacacaat aataatcgtc aatcaattca atcaatcaat 840
cggtgaatca atcataattg gatacgctaa aaactcttag cttagcttag gtctagaaat 900
atgaatatgt ctttcaacgt gtctgtggtt ttggtgtttt tggtgtttct gtgtggtttg 960
ctgtatggtg gctgagtggt tgagtggttg cgtttttagt tttcagtttt tctgtggctg 1020
tgactgtggg gtggtgcttt cattggcatg tcgcgtagcg ccaaacggcc cgtaaatatg 1080
caactatatg gcaaacaggc aagtaagtgt ttatatatat tgagagagag agatatatag 1140
agagttaagg aaatggctca atgcaattac aaccaattat ttacttgcag gactttgtag 1200
ttcttctcaa agggattacg gaggatgggt ttgggagggg gatttttacg gataattatt 1260
gtattttttt ttgaggagaa tgaaaatcta aatctaactc tagtcatctg ggtccaaccc 1320
gattagctct gcttgcaaac ggcccttagt cttacaacaa atggttaagc aaaatactct 1380
caaagcgata cggaattgcg gcgacatact ggtcatgatc agaattttcc aaaacgaatt 1440
```

caatatgccg atccagacga cggcgaagtt gataagagta ctttggataa catcgcccgt 1500
gaccacctcg tacagacgca acagaatcca tggcagccag gacacccaga atgcaaagac 1560
taatgcgaat gacatcatat gactagggtt cgataaattt tcagccagag cagttgcata 1620
ttctttatcg tgtatcggct cgccggacct aatcttacgc atcattacaa agatatagcc 1680
gtagttgtgt acgattgaaa tgaggctggg acctaagact aatatcgcca aggtggcgct 1740
atatgctgcc atattccccc agtccagcat gcaaatgtgc gtcatgttgt tctcgatggg 1800
catcgagtag cccaggattg gggacagca cagcagggcc gccgatatcc aagtgaacac 1860
catccagcat tggcatctgc acgataaagg tcacaagaga tcaaagatta atccagtttt 1920
aatccagatg aggaggaatt cggggttcca ttggcgctac tcacctcgtt ttcgtctgca 1980
ccgtttcata tctgagtggc ttccgcaccg caagatagcg atccactgat atccacatga 2040
aggtgtatac cgataccgcc cacagagtga cctccagata gccggtgaag cgacacacga 2100
tgtcgccgta catccattct ccggtcaaag cgggatacac ggagaagggc accaccagga 2160
gtccgcagag cagatcggcg atggccaggg acaagaggta gtagttgatc acctctgtgg 2220
gtcctgcgat gacagcaata tcgttagaaa acgaagtaga ggtgtattca aaaataatag 2280
gcaaatgcaa acaataatat ttaaatggat cattcaataa tttgggcgag tctatttgta 2340
cgacattcac tagaactatt tagaaaccta aggttcttaa atttatccaa taacaatcaa 2400
tttataggat aaaagtttag gtacaaaaag gggaaagaaa taatatgctt aagcacaatt 2460
ggtttttgaa tttttcaaa gcaattaaga caacaagaaa acaacacatt tgaatttttg 2520
gggaggatgc cgtgttactt cattcctttc actttcacat caaaatcatg gatgcatttt 2580
aagtataaac attggactta agtggtcaaa ataaattata ttcaacacat atgtaggtta 2640
cttatgaaaa tattatgctt gaggccaaac ccaaccgcga gattcgagat gttcctgtgg 2700
aaagtgttat gtgagtgtga tattcgtgca acgtaccctt gaagttggca taggtggcga 2760
taatgaggag gttggacagg acgatggcca cgcccagtat cgatatgagc accgccttgg 2820
tgagggcctc caccttggaa ttatcctgta ggtagctcat ttcctgcatt tttacgggct 2880
ggcactgtcg gttgaaatgg gcgctcctgt gctcgttgta tacgctccct tagctccttg 2940
tccgtttccc tcctccttgg ctcctcctgc tttcacctgc tttcgctgct gctttcaccc 3000
gctttctccg caccttcttg gcgggctgat tgttatgcaa aggccagtgg acgtggacgt 3060
gaccgtactt tacttccggt catggcggac gcccatgcaa ttgtcctgtg ccttcggttg 3120
tgtgtgtcgg tgtgtctgtg agtgtctatg tgtgcttagg tgagtgtgtc gatgtgtgag 3180
tgtgtgtgtg tgtctgggtg cgagtgttgg tggcgtcctt tattgtgcaa cgaggcttgc 3240
acacgcgaag gtcgtcggag tccttggcgc tgttttttcg tcctttctcc aacgttggcg 3300
ccaaatttgt ataaattttc tactgcctac ttttatttgg gctgcaactt ctactactgc 3360
ccacctttttg gcgccgagcc aagcatttga ttgttaccct cacaatcctg gcgtcgtggt 3420
taaggttaag gataaggttg tgctggtgtg tgtgtgtcta tgtgtgtgtc tgaagctcgt 3480
ttaaaatgct gcttgggcta aaatcttaca cgtaaaatct agtaattgcg tacaatttgg 3540
tttctatctt agaggatcgt taatgattta aagaattcgt tttgttttgc taaattcaca 3600
acgctctttg ctaagtttta atgaacgtca aaactgttag ttgtatgctt gaataattcg 3660
ctactagttt taatatcgat ttaaaaatgt acacatttca aactcgcaag tggaagaata 3720
atgattatta atttaactat ttaatgtgta ataattacgc ggttttcagg gctctttttc 3780
tttcggggtt agttgaaata tctttcgact ataattaaag cacagcacaa atggttttgt 3840
tgttttgtaa ggcttctgtt gtgctacatt taagtttgtt aaaatagttc gataatagtt 3900
cggttgttaa gtatcggttt tgctatttaa tttcgctgct tggtgattgt gattatcctt 3960
ttgcttgtaa attctggctc ttgttgctgt tgttggtgtt taaactccac gtcctgctgt 4020
cgttttaagc taattgtgcg catctgaaag gacaaaggat agcaattgta tgcctttagc 4080
gcttcataaa tagtttacag aatcttccat cttgagccaa tcaataaatt atcccacata 4140
cggcagcagg cggaaaaagc cttttagctg cacaaacgct ttaca 4185

<210> 140
<211> 1584
<212> DNA
<213> Drosophila

<400> 140
acacactcac acatcgacac actcacctaa gcacacatag acactcacag acacaccgac 60
acacacaacc gaaggcacag gacaattgca tgggcgtccg ccatgaccgg aagtaaagta 120
cggtcacgtc cacgtccact ggcctttgca taacaatcag cccgccaaga aggtgcggag 180
aaagcgggtg aaagcagcag cgaaagcagg tgaaagcagg aggagccaag gaggagggaa 240
acggacaagg agctaaggga gcgtatacaa cgagcacagg agcgcccatt tcaaccgaca 300
gtgccagccc gtaaaaatgc aggaaatgag ctacctacag gataattcca aggtggaggc 360
cctcaccaag gcggtgctca tatcgatact gggcgtggcc atcgtcctgt ccaacctcct 420

```
cattatcgcc acctatgcca acttcaaggg acccacagag gtgatcaact actacctctt 480
gtccctggcc atcgccgatc tgctctgcgg actcctggtg gtgcccttct ccgtgtatcc 540
cgctttgacc ggagaatgga tgtacggcga catcgtgtgt cgcttcaccg gctatctgga 600
ggtcactctg tgggcggtat cggtatacac cttcatgtgg atatcagtgg atcgctatct 660
tgcggtgcgg aagccactca gatatgaaac ggtgcagacg aaaacgagat gccaatgctg 720
gatggtgttc acttggatat cggcggccct gctgtgctgt cccccaatcc tgggctactc 780
gatgcccatc gagaacaaca tgacgcacat ttgcatgctg gactggggga atatggcagc 840
atatagcgcc accttggcga tattagtctt aggtcccagc ctcatttcaa tcgtacacaa 900
ctacggctat atctttgtaa tgatgcgtaa gattaggtcc ggcgagccga tacacgataa 960
agaatatgca actgctctgg ctgaaaattt atcgaaccct agtcatatga tgtcattcgc 1020
attagtcttt gcattctggg tgtcctggct gccatggatt ctgttgcgtc tgtacgaggt 1080
ggtcacgggc gatgttatcc aaagtactct tatcaacttc gccgtcgtct ggatcggcat 1140
attgaattcg ttttggaaaa ttctgatcat gaccagtatg tcgccgcaat tccgtatcgc 1200
tttgagagta ttttgcttaa ccatttgttg taagactaag ggccgtttgc aagcagagct 1260
aatcgggttg gacccagatg actagagtta gatttagatt ttcattctcc tcaaaaaaaa 1320
atacaataat tatccgtaaa aatccccctc ccaaacccat cctccgtaat ccctttgaga 1380
agaactacaa agtcctgcaa gtaaataatt ggttgtaatt gcattgagcc atttccttaa 1440
ctctctatat atctctctct ctcaatatat ataaacactt acttgcctgt ttgccatata 1500
gttgcatatt tacgggccgt ttggcgctac gcgacatgcc aatgaaagca ccaccccaca 1560
gtcacagcca cagaaaaact gaaa                                         1584
```

<210> 141
<211> 322
<212> PRT
<213> Drosophila

<400> 141

```
Met Gln Glu Met Ser Tyr Leu Gln Asp Asn Ser Lys Val Glu Ala Leu
 1               5                  10                  15
Thr Lys Ala Val Leu Ile Ser Ile Leu Gly Val Ala Ile Val Leu Ser
            20                  25                  30
Asn Leu Leu Ile Ile Ala Thr Tyr Ala Asn Phe Lys Gly Pro Thr Glu
        35                  40                  45
Val Ile Asn Tyr Tyr Leu Leu Ser Leu Ala Ile Ala Asp Leu Leu Cys
    50                  55                  60
Gly Leu Leu Val Val Pro Phe Ser Val Tyr Pro Ala Leu Thr Gly Glu
65                  70                  75                  80
Trp Met Tyr Gly Asp Ile Val Cys Arg Phe Thr Gly Tyr Leu Glu Val
                85                  90                  95
Thr Leu Trp Ala Val Ser Val Tyr Thr Phe Met Trp Ile Ser Val Asp
            100                 105                 110
Arg Tyr Leu Ala Val Arg Lys Pro Leu Arg Tyr Glu Thr Val Gln Thr
        115                 120                 125
Lys Thr Arg Cys Gln Cys Trp Met Val Phe Thr Trp Ile Ser Ala Ala
    130                 135                 140
Leu Leu Cys Cys Pro Pro Ile Leu Gly Tyr Ser Met Pro Ile Glu Asn
145                 150                 155                 160
Asn Met Thr His Ile Cys Met Leu Asp Trp Gly Asn Met Ala Ala Tyr
                165                 170                 175
Ser Ala Thr Leu Ala Ile Leu Val Leu Gly Pro Ser Leu Ile Ser Ile
            180                 185                 190
Val His Asn Tyr Gly Tyr Ile Phe Val Met Met Arg Lys Ile Arg Ser
        195                 200                 205
Gly Glu Pro Ile His Asp Lys Glu Tyr Ala Thr Ala Leu Ala Glu Asn
    210                 215                 220
Leu Ser Asn Pro Ser His Met Met Ser Phe Ala Leu Val Phe Ala Phe
225                 230                 235                 240
Trp Val Ser Trp Leu Pro Trp Ile Leu Leu Arg Leu Tyr Glu Val Val
                245                 250                 255
Thr Gly Asp Val Ile Gln Ser Thr Leu Ile Asn Phe Ala Val Val Trp
```

```
            260                 265                 270
Ile Gly Ile Leu Asn Ser Phe Trp Lys Ile Leu Ile Met Thr Ser Met
        275                 280                 285
Ser Pro Gln Phe Arg Ile Ala Leu Arg Val Phe Cys Leu Thr Ile Cys
    290                 295                 300
Cys Lys Thr Lys Gly Arg Leu Gln Ala Glu Leu Ile Gly Leu Asp Pro
305                 310                 315                 320
Asp Asp
```

<210> 142
<211> 4900
<212> DNA
<213> Drosophila

<400> 142

```
taaatatgca tattgaacga ggcaatgcaa tgcgattgag gtccggctgt gcgtttcagt 60
tgcgaaatga gttgcgaaca ggatttatat tcagatttag ccattagcca aatgcaacac 120
cttcgacgac gtgtggacac gtcgactttg attcggccaa ctgagtgatc cggggtagta 180
gtgtcagatg ccaacattac ttatgcactg aaagaaataa ataaaggatt tacatgtgta 240
gtaatgtata attgggctgc atgacatgca tgataatgat caactttcaa gcaataataa 300
ttctatgaca tcacgacatt catacggaca gaaagacgaa catggattta ccattaaaat 360
gtaattataa tataatataa taaagtatct aaagaatttt ttacttggat ctcctagcaa 420
ttttaaccat attctcaaat gcaaaatgtg aagttcttgc aaatagatac acagactaat 480
gggtccgcat ttgccaatta attgcctaca acgcccaggt tgctttgaat cacaggtgga 540
catttggaca ggtttgctca gctcactggc agctccactt tggccattca aaattgttta 600
tcgctttgac gagttgcagt agcaatttcg acaccgtatg cccataatga agaaacggca 660
ttgataataa caaagtaatg ataaattaga tgctgttgac gtctaattca tttttgatga 720
tagtattgat tttatacata aatgacacaa tttccctttt attttgtaaa acattatggc 780
attttcactt taatggtcct gtaaaacaca agttagttgt aaagcacaga aaaactttcc 840
atttcgatga caagaagtcc tttaaaattc aaataaatat ataattctaa ttctaaataa 900
ataattctaa gcaatggttt atgtatttaa actccttaaa ttttgttcca gtttttaact 960
agcacatata catcaaaaga gcggaactag atagtctaca atgatgcaag agaccggcaa 1020
tcaaatggga cagacccata tgcaccagcg ggttcctttc aacgacacgg ttctgaagga 1080
ctatcatctc acgagcaccg atatcgagaa gtttgtaaag ctatggcagg agtaccagat 1140
gaagaacatg acccctcagg tggatgagtg ccagggctat tgccaggggg aaatatacaa 1200
ctggttgcga gcctacaaca gcatccatgg atacgtttcg ctgatggtaa gtaatactcg 1260
tgattacaac tccccctggg cacgaaaccc gcaatgttgg ccaaaacaaa ggaatcggtt 1320
gactaatgat gagcaaggga agccaagacg atacacacag accttcgccg acagaccact 1380
gatgaatgta ggggacttct ttacctttga actcttgcca aagtcgtcct tttttcgaat 1440
taaagttata gcatgcgctc agcttaacca tcaaaatcac aaatgaaatt agtggccggg 1500
aatctcggtc ttttgtttat gtccgggttt cccttttcag ggctttttttg aaaaggatat 1560
aaattatatt agatgtagtt aagccgttga ggcgttaatg tgcttttttc attataattt 1620
tcagtcgtaa gtcattttgc gttgtagcag acggccttta ccacaggaat tgggtcagaa 1680
aataccgaga atccatttcc atcaaggcta gttcatagcc catcttttca aataaatata 1740
caataaatca aatacagttg aaaaatgtaa ctttattgct ttccttcgtt cgcacaaaaa 1800
atagtttacc tcttaaattg aaataaaatt ctgaaccaaa cttcggttac tgtagttcca 1860
tttttttttt tttttgcttg tgtgagcaat gttaagaaat gtaaatacga gaggtcacga 1920
cagttgagtt gctgctcacg agctgaataa ataaaatatt catccacttt ggtgtccaag 1980
gggagtgcat tcagttgaaa gtatgtgagc gctcacagtt tcgatttgac aaaggacgca 2040
ttttatgttg attctttctt gtttctttaa atcacctaga tttgcatttt cgggacaata 2100
gcaaacatct tgaacataat ggtcctgacc agaaaggaaa tggccaagac gcccataaat 2160
aatatactta agtggctggc ggtggccgat atgtttgtga tgctggaata tataccctat 2220
acatcgtatc agtatatata tatggggcca ggtaagcata catttagtaa taatatttaa 2280
taatagtctt ggttatagct ttacttgtta tttatttatt agtgttatca gctagattat 2340
tattataaag gtatactctt gatcaggatc tacacacgtc tgtctgtccg tttctaaggg 2400
tatattttgt taacaatttt gttaacaatt cccatatatc ctttcccctt tgcacttact 2460
tttagttgag taacgagtat gtggtagtcg aggcaatcgg ctatgacctt ctctgtaggt 2520
ttcctttctt ttttaatttt aagccgttta aattttttttg aattatttaa tattctttgt 2580
```

atggctgttg caggtgaaaa agatttgagc tacacctggg cggtttgtct gctggttcac 2640
atgcacttca cccagatcct gcatacgatt tccatcggat tgactgtcac gctggcggtg 2700
tggcggtatg tggctatcag gtgagatgat ccttcggcca gctatgcatg acttattccg 2760
gaattcgccg tcatccatgt agagtttgtc tccggttcac attcatgata atccacatcc 2820
gcaattcgaa cagacatccg gctaatctga ttaccacatc attgtggcaa tgtggcttct 2880
cttctcttcc cccgtctttt gtttatagcc aaatgttacc gtctctattc cacagacatc 2940
cgaacggggg ctgtgccaat ttcctgctag cacattcccg ggaggcgatc ctcctgccct 3000
tcatcctgtc gccgatcctc tgcctgccca cgtactttgt gttccaggtg cgtgagacgt 3060
acgacgtgga caaggtcaac tcggaggcca tgtatcatgt ttatttcgac aaggactcgg 3120
tgctctacag gtgcgcttag ataagaagtc tgatgtcatt tatgcatcct gcgaagaacc 3180
aattgtaaaa tatttcatcg tgaaaatcct taacagattc aacttctgga tacattccgt 3240
gctcatcaaa cttctgccgt gcggcatatt aattgtgatc agtgcagtac ttatgcacgt 3300
tttgtgcgag gcatcgagac gtcgcttgaa gctgagggat tacaacaatc ccgccaaata 3360
tgccatccag cttaatctga atgaaaccaa atccaagaag taagtacgca cttgactcgt 3420
ttagtgctta agaattcgcc tcatttgcag tttgccttgt tagtttcatg atctttgatt 3480
tttctgtttt tgaatggtct cggattttgt caaatctctt cttaggttcc tttttgccat 3540
cttttttata aatttgtctt tttgtttttg tatcttaggc cgcctcgctg tgatcgccgt 3600
aacgatcgga ctacccttttt gctggttgcg gtactggttt tgtttctgat cacggagttt 3660
ccgcagggat tacttggtct gctgtccgga gtgatggaga agtgcttctt cgcccactgc 3720
tatccgccgt ttggagagct gatggatctg ctggcactga tcaacgccgc cgtcggattc 3780
gtgctgtacg gtctgatgtc gaaacagttc cggaccacct tccgatcgct gttcatgaag 3840
cggcactttg gcagcaccga aatgaccaga ttgacccgcg taaccacgac ttgcgtctag 3900
ttgggcacat ctaaagatgt gatcatgggc aggagcaaac actagtgagg gatgaagtag 3960
gtgagggaca tccaagatgt tccggcccgg cattgaattt gaattttcta aagcataaac 4020
tttataaaag tttcgttgac taaacaaagc ctaaatatga gaaactgtga gtgtggaatt 4080
attgtacggg ctgaccagtt gatcagttga tttccacaat ggtttacttt taggtgtagt 4140
cctaaggcgt tgactggtgc atcgatcggt ttagaaattg attctgcttg tgatctggct 4200
gtttcgcaaa gtgccatttc atttccaaat aattacattg attagtttac cagttggtgt 4260
ttcaaattcg tattagtata taagtttcat aggctaagtc agctgtacaa agctgtacaa 4320
atacgagaca tttacttttc cataagatgc agaaaaaaaa acaggaatcc atcatcagca 4380
aaactgtagc catgctaaga ggactgaaga actcactgag ccccgagttg ctctctcttc 4440
ttagctggaa tcacgaaaga attctatttt gattgtatgg cgttaactgg ttttcaacaa 4500
agtgccaatg cattttaaat tggctaattg aataaccata ttttgagtac tttttaagtc 4560
gcttatacta tttataaata aatgtgaaca ctttcaaaaa aaatatgttt gatgtgagtt 4620
gcatacaaat gaaacgttta ttgctttcca ttaccgaaaa atgaacatct aattctgact 4680
gacaacagca ttcacctcca gtggcaattc gtccggcaga agttccgtgg cccgattgtg 4740
tcccaaatca gactcgccca tatccgcgga atccagcagg atgccattgc cctgattcga 4800
gatgcgcagc agggatttca gcttggtggc ttggaatttg tactgcagga taatgagatt 4860
gataaaatct gctctcgagg gagtttcgac cataatgtaa 4900

<210> 143
<211> 1245
<212> DNA
<213> Drosophila

<400> 143
atgatgcaag agaccggcaa tcaaatggga cagacccata tgcaccagcg ggttcctttc 60
aacgacacgg ttctgaagga ctatcatctc acgagcaccg atatcgagaa gtttgtaaag 120
ctatggcagg agtaccagat gaagaacatg acccctcagg tggatgagtg ccaggcctat 180
tgccaggggg aaatatacaa ctggttgcga gcctacaaca gcatccatgg atacgtttcg 240
ctgatgattt gcattttcgg gacaatagca aacatcttga acataatggt cctgaccaga 300
aaggaaatgg ccaagacgcc cataaataat atacttaagt ggctggcggt ggccgatatg 360
tttgtgatgc tggaatatat accctataca tcgtatcagt atatatatat ggggccaggt 420
gaaaaagatt tgagctacac ctgggcggtt tgtctgctgg ttcacatgca cttcacccag 480
atcctgcata cgatttccat cggattgact gtcacgctgg cggtgtggcg gtatgtggct 540
atcagacatc cgaacggggg ctgtgccaat ttcctgctag cacattcccg ggaggcgatc 600
ctcctgccct tcatcctgtc gccgatcctc tgcctgccca cgtactttgt gttccaggtg 660
cgtgagacgt acgacgtgga caaggtcaac tcggaggcca tgtatcatgt ttatttcgac 720
aaggactcgg tgctctacag attcaacttc tggatacatt ccgtgctcat caaacttctg 780
ccgtgcggca tattaattgt gatcagtgca gtacttatgc acgttttgtg cgaggcatcg 840

agacgtcgct tgaagctgag ggattacaac aatcccgcca aatatgccat ccagcttaat 900
ctgaatgaaa ccaaatccaa gaagccgcct cgctgtgatc gccgtaacga tcggactacc 960
cttttgctgg ttgcggtact ggttttgttt ctgatcacgg agtttccgca gggattactt 1020
ggtctgctgt ccggagtgat ggagaagtgc ttcttcgccc actgctatcc gccgtttgga 1080
gagctgatgg atctgctggc actgatcaac gccgccgtcg gattcgtgct gtacggtctg 1140
atgtcgaaac agttccggac caccttccga tcgctgttca tgaagcggca ctttggcagc 1200
accgaaatga ccagattgac ccgcgtaacc acgacttgcg tctag 1245

<210> 144
<211> 414
<212> PRT
<213> Drosophila

<400> 144
Met Met Gln Glu Thr Gly Asn Gln Met Gly Gln Thr His Met His Gln
1 5 10 15
Arg Val Pro Phe Asn Asp Thr Val Leu Lys Asp Tyr His Leu Thr Ser
20 25 30
Thr Asp Ile Glu Lys Phe Val Lys Leu Trp Gln Glu Tyr Gln Met Lys
35 40 45
Asn Met Thr Pro Gln Val Asp Glu Cys Gln Gly Tyr Cys Gln Gly Glu
50 55 60
Ile Tyr Asn Trp Leu Arg Ala Tyr Asn Ser Ile His Gly Tyr Val Ser
65 70 75 80
Leu Met Ile Cys Ile Phe Gly Thr Ile Ala Asn Ile Leu Asn Ile Met
85 90 95
Val Leu Thr Arg Lys Glu Met Ala Lys Thr Pro Ile Asn Asn Ile Leu
100 105 110
Lys Trp Leu Ala Val Ala Asp Met Phe Val Met Leu Glu Tyr Ile Pro
115 120 125
Tyr Thr Ser Tyr Gln Tyr Ile Tyr Met Gly Pro Gly Glu Lys Asp Leu
130 135 140
Ser Tyr Thr Trp Ala Val Cys Leu Leu Val His Met His Phe Thr Gln
145 150 155 160
Ile Leu His Thr Ile Ser Ile Gly Leu Thr Val Thr Leu Ala Val Trp
165 170 175
Arg Tyr Val Ala Ile Arg His Pro Asn Gly Gly Cys Ala Asn Phe Leu
180 185 190
Leu Ala His Ser Arg Glu Ala Ile Leu Leu Pro Phe Ile Leu Ser Pro
195 200 205
Ile Leu Cys Leu Pro Thr Tyr Phe Val Phe Gln Val Arg Glu Thr Tyr
210 215 220
Asp Val Asp Lys Val Asn Ser Glu Ala Met Tyr His Val Tyr Phe Asp
225 230 235 240
Lys Asp Ser Val Leu Tyr Arg Phe Asn Phe Trp Ile His Ser Val Leu
245 250 255
Ile Lys Leu Leu Pro Cys Gly Ile Leu Ile Val Ile Ser Ala Val Leu
260 265 270
Met His Val Leu Cys Glu Ala Ser Arg Arg Arg Leu Lys Leu Arg Asp
275 280 285
Tyr Asn Asn Pro Ala Lys Tyr Ala Ile Gln Leu Asn Leu Asn Glu Thr
290 295 300
Lys Ser Lys Lys Pro Pro Arg Cys Asp Arg Arg Asn Asp Arg Thr Thr
305 310 315 320
Leu Leu Leu Val Ala Val Leu Val Leu Phe Leu Ile Thr Glu Phe Pro
325 330 335
Gln Gly Leu Leu Gly Leu Leu Ser Gly Val Met Glu Lys Cys Phe Phe
340 345 350
Ala His Cys Tyr Pro Pro Phe Gly Glu Leu Met Asp Leu Leu Ala Leu
355 360 365

```
Ile Asn Ala Ala Val Gly Phe Val Leu Tyr Gly Leu Met Ser Lys Gln
    370                 375                 380
Phe Arg Thr Thr Phe Arg Ser Leu Phe Met Lys Arg His Phe Gly Ser
385                 390                 395                 400
Thr Glu Met Thr Arg Leu Thr Arg Val Thr Thr Thr Cys Val
                405                 410


<210> 145
<211> 4116
<212> DNA
<213> Drosophila

<400> 145
gttaaatgat ttacaaacaa gccgacatct cgggaaccat taaagagaca tttaaagtca 60
aaaagtggcc aactgaaacc cgttaaaatc cgcttgcgac cgcaacttgg cccgactaag 120
gcaaaaaata acataatcct gtcaagtcat gcagacaggg cgatttacat tataatttta 180
aatggcaatg gattagtcat catcagtacc tgcttatctg gttttcctga tttttaaatc 240
aaatctgtga ctgtatctct tacctttcct cttgtacctc gttattataa aatgcgaaaa 300
ccctctgcca aaggacacta taaaacgaat tttttaaaag acgccaaaca aaatggcact 360
ttttgttgag aagaaaatat gccgataaga acgttttact tgtaagacgc tgcagaaaga 420
ttaggaacga aaactatcga caatgtttac ttgtataatc tgagctttct aaaacacaaa 480
aatagtgtac ttttaccata cgcttaattt cgaaattaat ccgggatata cagttgggga 540
tccagatttt aaattcgtga acacagctgc tggagtaggg catttcctga tcgatctgtt 600
agttttcgga cttttacttt ggttgttgta aacaatatcc gcaagtttct cgcaaagtgg 660
gttgcgttgt gccgtgcctc agtagagcac tacacactac cttcgattga gttcgctggg 720
atttggattt cgttcacagt ccaagtttgg accgtgctag gagaagataa tgatccgatc 780
actcaattaa cttcatgcac ccagtgacgt catcagtgct ttactattcg gtgtctatat 840
agtctatttt gtttaatttc ttacacccta catttcgata cgattcgatt ttcgtttgga 900
aattgtctga aggtaaacaa gcataaattg gacgttcaag cgtaaacaaa tcggttttcg 960
gagcaaaata caaaaatcgt cagttgagaa agtctctcaa atggcgcagt tctgcattct 1020
tggagtcctc ctcattcttt ctggaacaca ctgctcgtgg ggattccatg aagagacaca 1080
ttatccgtgt gcctttatcg atacggcaaa cataacgggg agctacggct tagatggccc 1140
atttgtgcac aactggacgg tgattccccg tcacttcgta gccgtgtacg actttgtcat 1200
cgaaaatggc atccgcatcc ccgcaagcag gcatctcaga gcctgtgtct gtaagaccaa 1260
accctgtgta cggatttgct gtttacgggg agaaatctac gacctggaga agaggcagtg 1320
cttagtccct gtggccgggg tgtccagcct tccctcccat agccatatgg aggtggagtt 1380
gggcaacggc agtctgcgtc tagtgaaact gcaaccccgt tttagcatcc acgtggaaac 1440
accatgtgag cacatgaaag ccgtaaccaa gggctcggaa tacgttcact ggacacttca 1500
tgaggtcggt tgtgaaataa accttccatc acattaatgg ctttacatcg ccttttcaga 1560
acggaaccat cagccaccga ggacacatat tttctaagca ttactgcttc actccgctgc 1620
tccatgggaa ctccacctgg gagtggcagc cattggcctg tgcaccagaa aaactatatt 1680
tcgtcttggg ggtgcgggag tggacatacg ccatatgtaa gtagcaagag gaaacatttt 1740
tccagcaact ttcagcaatt cattatccca aatatgcaca atactgtttt cacaacaata 1800
cataaaagta attgctaaag caccaaaaaa caatcgatga tagcattaat agtttgatag 1860
tcagaattta atttacccag cgtattcact ttccaggtct cttgatagct attctatcaa 1920
tgtttatcgt gctgatggtc taccttatgt gttcggaaat gcgcaacagt ttctacggtg 1980
tggccataaa ggcgtacgcc atttgtatga tcttgggata tgccctactg gcttatctca 2040
cgctccacaa cccagcaaat ctctcaaatg cggcttgtcg tattctacgt aagtggtaaa 2100
tttgagtatt ctgtgatatt taaacattcc cttctttttt ttttttagca agtttggcgt 2160
tgatgaactt ggtgctttcc ttttacattc tgagctttat tgcatttaag ctgtacttga 2220
gcttttatgg cgttgtcttc acaaaactga tgttttggct aatatttact cccatcgtac 2280
tggtagccgt gggctggtct ttctttgtgg gtttcagtta ttacggctct aggctcatct 2340
tcggaggcga tacatgctgg tttgatcgta agcactgcca atccatatat gttttataat 2400
aattaattgg tattccacat cgctagcacg caattggtcc gttatgatat atttttacgc 2460
tcctgttttc gtggcttgcg ccattagcgg tttcttctat gtcctgagcc agatctacat 2520
ccgtgatcag ccagatatcg agacggagaa gagcttcgag tccatagaga agaaccgatt 2580
caaatcattt tggaagtact ttggctacac ggcagtcgta tgggtcgttt gcatatgctc 2640
ctttgccttt aactactact gggagaacag atcccacctt aactacgctg taagcttctg 2700
catggccttc cacgggtttg cggccctcta tgcgctgatt ggaaagaatc agcaaatcca 2760
```

aaattttctg cggcgcatag ataagtaagt tgataactat gaattttaag tactactgtt 2820
aatatatact gttgcagtgg cgaagatacc tgtgaaaact cagttccgct ctcgagtttt 2880
ggttaacttt tcttgttgtg cccactgtag cgttagtgta actggaagca gttgagtgta 2940
tcgacaatat ctgtgcaatg tttttaaaat cacttatcag catagacaaa aaccatgtat 3000
tttaaatgta aaattttgta ggatgtgtgt aaaatatatt tgtaaaattt gaactcgaac 3060
aaaaaagtac acttctaccc tttaaacgaa attacgataa taaagataca aaaaataaat 3120
atatctgaaa aaaaaatcag gtctggaatt tctacaattg ttaaatatga taacaagtcc 3180
caacaatcct ttcagtatag aacaaaacat ggaatatact taagccatgt gaaagaacac 3240
ataacatatt aatattctaa acaatttgag cagttttctt gagtttacgc ctccaacaat 3300
ttatccttca cctttcttga gttccaaatt cctggaccgt gatcaagttt ctccagtctt 3360
atttcaagtc ccgattttac tgtaccctat ccgcatcgcc aagcccggtg ttacttcccc 3420
gatctctcta tattataaat gtgcgccgat tttatttaac atcagaggtc aatggccgca 3480
aaacagacag agagagagag agagaaagtt ggtcagagcc acgagaactg catcccgaca 3540
gggggctgtc aagagcaagg aatcaagtgg aatgttggtg gtgctctttt gaaaaggctt 3600
cgcaaacttt ccccctgtac accccttgc caaggcctaa tcgcagaaaa cccacccggg 3660
tcggagtccc caattcattt tcatgccgat tatgcttcgg gttttggtcc tgctcttttt 3720
ggctgggcgt acagcagcag gactcggggt acacactaag agaggtctaa gaatatagag 3780
atagacaagg aaagcagatg gatcgataaa acgtagtaaa aatgctcgta gtaatggaag 3840
ctggaaataa catggcacaa acaggcccca aactgcatcg acaacaccat cacaatccag 3900
acgtgcgcac tggaaaatcc cgagaggtag gacatgtttg gcgcgttttt cgactgcatg 3960
tgtgcgagtg ttctgttttg gttagatacg tatgaaaaag tgaaaaagcc tatattgacc 4020
gcgcccggta gatcgcatag gcaatatcga tatgggaaat cattgactaa catttgcaga 4080
catttagttt gggacacatt aaagttaatt tgactc 4116

<210> 146
<211> 1709
<212> DNA
<213> Drosophila

<400> 146
atggcgcagt tctgcattct tggagtcctc ctcattcttt ctggaacaca ctgctcgtgg 60
ggattccatg aagagacaca ttatccgtgt gcctttatcg atacggcaaa cataacgggg 120
agctacggct tagatggccc atttgtgcac aactggacgg tgattccccg tcacttcgta 180
gccgtgtacg actttgtcat cgaaaatggc atccgcatcc ccgcaagcag gcatctcaga 240
gcctgtgtct gtaagaccaa accctgtgta cggatttgct gtttacgggg agaaatctac 300
gacctggaga agaggcagtg cttagtccct gtggccgggg tgtccagcct tccctcccat 360
agccatatgg aggtggagtt gggcaacggc agtctgcgtc tagtgaaact gcaaccccgt 420
tttagcatcc acgtggaaac accatgtgag cacatgaaag ccgtaaccaa gggctcggaa 480
tacgttcact ggacacttca tgagaacgga accatcagcc accgaggaca catattttct 540
aagcattact gcttcactcc gctgctccat gggaactcca cctgggagtg gcagccattg 600
gcctgtgcac cagaaaaact atatttcgtc ttgggggtgc gggagtggac atacgccata 660
tgtctcttga tagctattct atcaatgttt atcgtgctga tggtctacct tatgtgttcg 720
gaaatgcgca acagtttcta cggtgtggcc ataaaggcgt acgccatttg tatgatcttg 780
ggatatgccc tactggctta tctcacgctc cacaacccag caaatctctc aaatgcggct 840
tgtcgtattc taccaagttt ggcgttgatg aacttggtgc tttcctttta cattctgagc 900
tttattgcat ttaagctgta cttgagcttt tatggcgttg tcttcacaaa actgatgttt 960
tggctaatat ttactcccat cgtactggta gccgtgggct ggtctttctt tgtgggtttc 1020
agttattacg gctctaggct catcttcgga ggcgatacat gctggtttga tccacgcaat 1080
tggtccgtta tgatatattt ttacgctcct gttttcgtgg cttgcgccat tagcggtttc 1140
ttctatgtcc tgagccagat ctacatccgt gatcagccag atatcgagac ggagaagagc 1200
ttcgagtcca tagagaagaa ccgattcaaa tcattttgga agtactttgg ctacacggca 1260
gtcgtatggg tcgtttgcat atgctccttt gcctttaact actactggga gaacagatcc 1320
caccttaact acgctgtaag cttctgcatg gccttccacg ggtttgcggc cctctatgcg 1380
ctgattggaa agaatcagca aatccaaaat tttctgcggc gcatagataa tggcgaagat 1440
acctgtgaaa actcagttcc gctctcgagt tttggttaac ttttcttgtt gtgcccactg 1500
tagcgttagt gtaactggaa gcagttgagt gtatcgacaa tatctgtgca atgtttttaa 1560
aatcacttat cagcatagac aaaaaccatg tattttaaat gtaaaatttt gtaggatgtg 1620
tgtaaaatat atttgtaaaa tttgaactcg aacaaaaaag tacacttcta ccctttaaac 1680
gaaattacga taataaagat acaaaaaat 1709

<210> 147
<211> 492
<212> PRT
<213> Drosophila

<400> 147
Met Ala Gln Phe Cys Ile Leu Gly Val Leu Leu Ile Leu Ser Gly Thr
1 5 10 15
His Cys Ser Trp Gly Phe His Glu Glu Thr His Tyr Pro Cys Ala Phe
20 25 30
Ile Asp Thr Ala Asn Ile Thr Gly Ser Tyr Gly Leu Asp Gly Pro Phe
35 40 45
Val His Asn Trp Thr Val Ile Pro Arg His Phe Val Ala Val Tyr Asp
50 55 60
Phe Val Ile Glu Asn Gly Ile Arg Ile Pro Ala Ser Arg His Leu Arg
65 70 75 80
Ala Cys Val Cys Lys Thr Lys Pro Cys Val Arg Ile Cys Cys Leu Arg
85 90 95
Gly Glu Ile Tyr Asp Leu Glu Lys Arg Gln Cys Leu Val Pro Val Ala
100 105 110
Gly Val Ser Ser Leu Pro Ser His Ser His Met Glu Val Glu Leu Gly
115 120 125
Asn Gly Ser Leu Arg Leu Val Lys Leu Gln Pro Arg Phe Ser Ile His
130 135 140
Val Glu Thr Pro Cys Glu His Met Lys Ala Val Thr Lys Gly Ser Glu
145 150 155 160
Tyr Val His Trp Thr Leu His Glu Asn Gly Thr Ile Ser His Arg Gly
165 170 175
His Ile Phe Ser Lys His Tyr Cys Phe Thr Pro Leu Leu His Gly Asn
180 185 190
Ser Thr Trp Glu Trp Gln Pro Leu Ala Cys Ala Pro Glu Lys Leu Tyr
195 200 205
Phe Val Leu Gly Val Arg Glu Trp Thr Tyr Ala Ile Cys Leu Leu Ile
210 215 220
Ala Ile Leu Ser Met Phe Ile Val Leu Met Val Tyr Leu Met Cys Ser
225 230 235 240
Glu Met Arg Asn Ser Phe Tyr Gly Val Ala Ile Lys Ala Tyr Ala Ile
245 250 255
Cys Met Ile Leu Gly Tyr Ala Leu Leu Ala Tyr Leu Thr Leu His Asn
260 265 270
Pro Ala Asn Leu Ser Asn Ala Ala Cys Arg Ile Leu Pro Ser Leu Ala
275 280 285
Leu Met Asn Leu Val Leu Ser Phe Tyr Ile Leu Ser Phe Ile Ala Phe
290 295 300
Lys Leu Tyr Leu Ser Phe Tyr Gly Val Val Phe Thr Lys Leu Met Phe
305 310 315 320
Trp Leu Ile Phe Thr Pro Ile Val Leu Val Ala Val Gly Trp Ser Phe
325 330 335
Phe Val Gly Phe Ser Tyr Tyr Gly Ser Arg Leu Ile Phe Gly Gly Asp
340 345 350
Thr Cys Trp Phe Asp Pro Arg Asn Trp Ser Val Met Ile Tyr Phe Tyr
355 360 365
Ala Pro Val Phe Val Ala Cys Ala Ile Ser Gly Phe Phe Tyr Val Leu
370 375 380
Ser Gln Ile Tyr Ile Arg Asp Gln Pro Asp Ile Glu Thr Glu Lys Ser
385 390 395 400
Phe Glu Ser Ile Glu Lys Asn Arg Phe Lys Ser Phe Trp Lys Tyr Phe
405 410 415
Gly Tyr Thr Ala Val Val Trp Val Val Cys Ile Cys Ser Phe Ala Phe
420 425 430

```
Asn Tyr Tyr Trp Glu Asn Arg Ser His Leu Asn Tyr Ala Val Ser Phe
        435                 440                 445
Cys Met Ala Phe His Gly Phe Ala Ala Leu Tyr Ala Leu Ile Gly Lys
    450                 455                 460
Asn Gln Gln Ile Gln Asn Phe Leu Arg Arg Ile Asp Asn Gly Glu Asp
465                 470                 475                 480
Thr Cys Glu Asn Ser Val Pro Leu Ser Ser Phe Gly
                485                 490
```

```
<210> 148
<211> 3772
<212> DNA
<213> Drosophila

<400> 148
agtcagtgcc tggacggatc gcacgagtcc ttctactaag cgctaccagt ggtggtcgta 60
atccccccaat ccaaacccat acccaaagct tcgctgcatt ttcgaataca aacactgaac 120
tatccgtcgt tgtaaagcta ccttatccac tacgattact acactatcct acaaaaaact 180
atgctggtta aaagactata cacctagctc catatactat atatatatat acaatatcga 240
ggcatgtgtt gtgtgggcga gtgagtgtgt gcgtgtgact tctgattgag ttgccatttg 300
attccaagta ttttcgagtg cttagttttg taatttgtag ctagccgtta atgtttgtat 360
aaactgtttc attggtttcc cactccgctt tattcgataa gctcgttttc tagattttag 420
ctttagcact cttagtttaa tgtttaaagt tagtttcgct gttaagttta caccgtagtc 480
gaggcgtctt gagtagttct tttgtacatt aagcttacac ttagctctac attagatata 540
ccatgtcgaa cctacgagtc tacgaacata cgaacatata tatgtgcgtt gaaagcgttg 600
tattttttga agccgaactt atgcaaccga aaggagacaa taaacaaatg tgctataaag 660
gacgaaagtg tcgcaaataa atgtgagcca ctgctaaaga agaccgattt gcgcaaggtg 720
ttttactcgg cggcggcggc aagaagcgtt tggggcacaa acgacggatg agtgattgcc 780
actcgtacca tcatcatgac gatggcgtgt ggaacgacgt gcgcatagta atgacagcgg 840
atggcaatct gaacgataag cggcacccga taatgaaaca ttaaactaaa ccgtcggtag 900
tcgctggcca taaatctgtc ggcagtctgc tcgcagaagc ttttaggtgc ggaccccttg 960
cgcggcatac tttcgggctt ccgttccgtt ccgttcgcca atggccctgc tccactacac 1020
cttcgaccag ctcgagttgt acctagaatg ggcgttcgcc cagcacggag aggccacgcc 1080
gattccctcg atccagccgt atccgggcgt gttcgtgggc gatctcagcc agctgaaccg 1140
cttcaagcga catgccttct ccgccgtcgt gggcacgctc ttcgtcctcg ccttctgcgg 1200
caacctgagc accctctatg tgaacagccg ccggaagctg cggccccttct tccgggcctg 1260
cctcatctca ctggcctgca gtgaccttgt ctccagcata ttctgcaccg tatcctacat 1320
ggcccagttc caggcgcagt atctgcagct ctgggtgagg cggcagggat tgagcagact 1380
caacacaagc taattacaaa tctccttaga caattggtgg cttcatgtgc aagtttgtgc 1440
cgttcataac caccacgtcg gtgttgtccg gcagcctgac cctggtggcc atcgccttgg 1500
accgctactt ggccgtgatg cgacccgtgc tggggttctg gagtcccgac aaacgcttca 1560
gcactctgag catgctgctc atctgggcat gttcgatagg atcctctggc ccgctgctgg 1620
gcatctacga ctaccgaaag atctatctcc tggatgtgga ggactccagc gaggagagcg 1680
aggaggtcgt gaccgctgtg cccgaggagc tggtggtcac cgagctggaa atggttcaca 1740
tgtgcctggc cggggacgta agtcaacgac tggattatcg cttaatacag ttgacagcga 1800
ttaaggggat ttggacgggg ggatgtcaag ttgttgaatc aatgttccga ttatcttatc 1860
agccgcgtcg aaaacaattt atgcaattct ataaatattc cgttttagca cgacgtcgga 1920
ctctactacg tcatactgtt caccctgatc ttcctgccgt gcattgtgtc cttcctatgg 1980
ctgaatgccg tgattgcgag gcagctctgg ctgcggagac actaccacca ggagcagcag 2040
gagcagcatc aggagcccaa ggagggtcag tttaagacca tggcgaatgg cggggatctc 2100
ttgatgccct ccaccttggt cagtgccatg ggagtggccg ttccctttgc gctggacaac 2160
actcccttgc cgcccaagag taccgtcaac gctcctggca agaagaccac tgctgctgct 2220
ctagccaggg aggccaggca ccgcaagatg gtggtggtgg tgctgctaat gatggcggtt 2280
ttcatatgcc ttcgactgcc cgcctgggtc ttcctgatca tgcggcttta tggatcctac 2340
tccgagccca tcgactggct cctgtacttc agcttcggca tactgaacct tttcagttgc 2400
gccctaaatc ccatcttcta caccttcctc acccagacca ttaggacctt gacgctggtg 2460
aaacacaaaa ttcagggatt tcttggctgt ccgcctggta aagtgccgga tggtatgccc 2520
acggatcaga tggacaaaag cgggtgctgc tgcggcctga ggccacccac cttcacatgg 2580
aggtgccatc ccagccgcga tagggcagca gccaccgtaa tccgggatgt ggaccagcca 2640
```

```
gatccgccct ccgaccaggt gcaaccggat ccgagtagcc tacggcggtt cctcagctac 2700
aagcaggagg tgttcaccat ttacaagcag tgcggtgact cgtcgagtgc gtcaatcgag 2760
tcctccgcct gataggcggt ggtcataaat cacccgggtg catttgtcgg atacgggggc 2820
caacggttag cagaacccca gaacttgatt gatggatgga tccagccccc attcactctg 2880
tacaggcgca aaaacggttt gtcacggtgg gtatatcacc cacgaataca ctagcacgca 2940
tataagatat aagctcacat attatgtacc tcatcaaatg tgtaaatatg aggaataaac 3000
gaatgggtgt tggcagctaa ctaaccatta acacatggcg caatcgaaat cgggaaggtc 3060
tcagctatta ttgtgacgaa aacctaagac ctttcaagta ccacttaaaa ctatatttag 3120
agaacttaag caaattgtga gattaaaatg caatgcaact cctaaaggtg cgaatatacc 3180
tattcttaaa tatttttact ctagaaatac tgacatcttt atgtgcttag ttttcgttca 3240
gacagactat ctttgttatt tctaagaacg tagtaattag ttaatggaat taattagtaa 3300
ccgaaactta aaattgagtc atgccaaaaa aagagaattg gcctgacaag acacacatta 3360
tagagtcagc agatcgtgga cgagacaatg agcccgcaat gcaatggaac tagattgtat 3420
catatactag atacttccgc ttggaacaat ggttaacaga tctcgtgaga gatggggctg 3480
aaagcagttg ttaattgtca aatgttgagc acgccagtta aatttagcta tgcagcgaaa 3540
gctctctccc atactctctt cccatgaact tttttaagcc gcagagcccg aagctcgact 3600
caaagctgaa ttctgcgttt tctggcattc agtgtcggcg gtttgcggcg agcggttgga 3660
ttcggattct ttggtcggta aaacgcgaaa tcttctgtgc ggcgggtgaa atcttcggca 3720
gaagaaacaa caaagcgaca agtgtggagt gtggagtgca taccgtacca cg         3772
```

<210> 149
<211> 1566
<212> DNA
<213> Drosophila

<400> 149

```
atggccctgc tccactacac cttcgaccag ctcgagttgt acctagaatg ggcgttcgcc 60
cagcacggag aggccacgcc gattccctcg atccagccgt atccgggcgt gttcgtgggc 120
gatctcagcc agctgaaccg cttcaagcga catgccttct ccgccgtcgt gggcacgctc 180
ttcgtcctcg ccttctgcgg caacctgagc accctctatg tgaacagccg ccggaagctg 240
cggcccttct tccgggcctg cctcatctca ctggcctgca gtgaccttgt ctccagcata 300
ttctgcaccg tatcctacat ggcccagttc caggcgcagt atctgcagct ctggacaatt 360
ggtggcttca tgtgcaagtt tgtgccgttc ataaccacca cgtcggtgtt gtccggcagc 420
ctgaccctgg tggccatcgc cttggaccgc tacttggccg tgatgcgacc cgtgctgggg 480
ttctggagtc ccgacaaacg cttcagcact ctgagcatgc tgctcatctg ggcatgttcg 540
ataggatcct ctggcccgct gctgggcatc tacgactacc gaaagatcta tctcctggat 600
gtggaggact ccagcgagga gagcgaggag gtcgtgaccg ctgtgcccga ggagctggtg 660
gtcaccgagc tggaaatggt tcacatgtgc ctggccgggg accacgacgt cggactctac 720
tacgtcatac tgttcaccct gatcttcctg ccgtgcattg tgtccttcct atggctgaat 780
gccgtgattg cgaggcagct ctggctgcgg agacactacc accaggagca gcaggagcag 840
catcaggagc ccaaggaggg tcagtttaag accatggcga atggcgggga tctcttgatg 900
ccctccacct tggtcagtgc catgggagtg gccgttccct ttgcgctgga caacactccc 960
ttgccgccca agagtaccgt caacgctcct ggcaagaaga ccactgctgc tgctctagcc 1020
agggaggcca ggcaccgcaa gatggtggtg gtggtgctgc taatgatggc ggttttcata 1080
tgccttcgac tgcccgcctg ggtcttcctg atcatgcggc tttatggatc ctactccgag 1140
cccatcgact ggctcctgta cttcagcttc ggcatactga accttttcag ttgcgcccta 1200
aatcccatct tctacacctt cctcacccag accattagga ccttgacgct ggtgaaacac 1260
aaaattcagg gatttcttgg ctgtccgcct ggtaaagtgc cggatggtat gcccacggat 1320
cagatggaca aaagcgggtg ctgctgcggc ctgaggccac ccaccttcac atggaggtgc 1380
catcccagcc gcgatagggc agcagccacc gtaatccggg atgtggacca gccagatccg 1440
ccctccgacc aggtgcaacc ggatccgagt agcctacggc ggttcctcag ctacaagcag 1500
gaggtgttca ccatttacaa gcagtgcggt gactcgtcga gtgcgtcaat cgagtcctcc 1560
gcctga                                                            1566
```

<210> 150
<211> 521
<212> PRT
<213> Drosophila

<400> 150

```
Met Ala Leu Leu His Tyr Thr Phe Asp Gln Leu Glu Leu Tyr Leu Glu
 1               5                  10                  15
Trp Ala Phe Ala Gln His Gly Glu Ala Thr Pro Ile Pro Ser Ile Gln
            20                  25                  30
Pro Tyr Pro Gly Val Phe Val Gly Asp Leu Ser Gln Leu Asn Arg Phe
        35                  40                  45
Lys Arg His Ala Phe Ser Ala Val Val Gly Thr Leu Phe Val Leu Ala
    50                  55                  60
Phe Cys Gly Asn Leu Ser Thr Leu Tyr Val Asn Ser Arg Arg Lys Leu
65                  70                  75                  80
Arg Pro Phe Phe Arg Ala Cys Leu Ile Ser Leu Ala Cys Ser Asp Leu
                85                  90                  95
Val Ser Ser Ile Phe Cys Thr Val Ser Tyr Met Ala Gln Phe Gln Ala
            100                 105                 110
Gln Tyr Leu Gln Leu Trp Thr Ile Gly Gly Phe Met Cys Lys Phe Val
        115                 120                 125
Pro Phe Ile Thr Thr Thr Ser Val Leu Ser Gly Ser Leu Thr Leu Val
    130                 135                 140
Ala Ile Ala Leu Asp Arg Tyr Leu Ala Val Met Arg Pro Val Leu Gly
145                 150                 155                 160
Phe Trp Ser Pro Asp Lys Arg Phe Ser Thr Leu Ser Met Leu Leu Ile
                165                 170                 175
Trp Ala Cys Ser Ile Gly Ser Ser Gly Pro Leu Leu Gly Ile Tyr Asp
            180                 185                 190
Tyr Arg Lys Ile Tyr Leu Leu Asp Val Glu Asp Ser Ser Glu Glu Ser
        195                 200                 205
Glu Glu Val Val Thr Ala Val Pro Glu Glu Leu Val Val Thr Glu Leu
    210                 215                 220
Glu Met Val His Met Cys Leu Ala Gly Asp His Asp Val Gly Leu Tyr
225                 230                 235                 240
Tyr Val Ile Leu Phe Thr Leu Ile Phe Leu Pro Cys Ile Val Ser Phe
                245                 250                 255
Leu Trp Leu Asn Ala Val Ile Ala Arg Gln Leu Trp Leu Arg Arg His
            260                 265                 270
Tyr His Gln Glu Gln Gln Glu Gln His Gln Glu Pro Lys Glu Gly Gln
        275                 280                 285
Phe Lys Thr Met Ala Asn Gly Gly Asp Leu Leu Met Pro Ser Thr Leu
    290                 295                 300
Val Ser Ala Met Gly Val Ala Val Pro Phe Ala Leu Asp Asn Thr Pro
305                 310                 315                 320
Leu Pro Pro Lys Ser Thr Val Asn Ala Pro Gly Lys Lys Thr Thr Ala
                325                 330                 335
Ala Ala Leu Ala Arg Glu Ala Arg His Arg Lys Met Val Val Val Val
            340                 345                 350
Leu Leu Met Met Ala Val Phe Ile Cys Leu Arg Leu Pro Ala Trp Val
        355                 360                 365
Phe Leu Ile Met Arg Leu Tyr Gly Ser Tyr Ser Glu Pro Ile Asp Trp
    370                 375                 380
Leu Leu Tyr Phe Ser Phe Gly Ile Leu Asn Leu Phe Ser Cys Ala Leu
385                 390                 395                 400
Asn Pro Ile Phe Tyr Thr Phe Leu Thr Gln Thr Ile Arg Thr Leu Thr
                405                 410                 415
Leu Val Lys His Lys Ile Gln Gly Phe Leu Gly Cys Pro Pro Gly Lys
            420                 425                 430
Val Pro Asp Gly Met Pro Thr Asp Gln Met Asp Lys Ser Gly Cys Cys
        435                 440                 445
Cys Gly Leu Arg Pro Pro Thr Phe Thr Trp Arg Cys His Pro Ser Arg
    450                 455                 460
Asp Arg Ala Ala Ala Thr Val Ile Arg Asp Val Asp Gln Pro Asp Pro
465                 470                 475                 480
```

```
Pro Ser Asp Gln Val Gln Pro Asp Pro Ser Ser Leu Arg Arg Phe Leu
                485                 490                 495
Ser Tyr Lys Gln Glu Val Phe Thr Ile Tyr Lys Gln Cys Gly Asp Ser
            500                 505                 510
Ser Ser Ala Ser Ile Glu Ser Ser Ala
        515                 520


<210> 151
<211> 8599
<212> DNA
<213> Drosophila

<400> 151
aggacacatt ggcgtactta gtctagacca ctgggtgagc ttaatgtttt gccgttcggt 60
tatatatcat gcggtgattt gtggacttag cataaaacct aactgaacgg cccggcgaca 120
tcccgactct ctggcgactt tggacgggcc acgataaatc aaggcgcaga tccgggagcc 180
cggcatatag tggatatagt atgcgtggcg aaacttaata agagatgcga caacaaacaa 240
atgatacaga tagacagacg gcaggcgaca gtcgcggtgg tgcgagaggt gggtaggtgg 300
gtaggtggga aatgggattg tggaaagcgg aaaaatcttt aaagcaatgc cgctgcacag 360
gacatttttc accgcgccat ccattcgcag acggatcgcc aaagagagac atcaacgtga 420
ctcgcgggga gaacaagtgt tttgatttat cgccagcaat ggcaatggca gggcggacgg 480
ggattgccat ccaattatac tgctaacttt cattatgtca ctgtgctatt tcgggaccgg 540
attcgccgcc atattggccc attgttgcgc cattgtctat gcacagctct ctgtttagca 600
gtgacacact gcggtttcgg tggcaggact catctccatc ctgccgcctg ccttagtcgt 660
cccttgccgc ttcctcggtg gactccgtct gataaatcgt cagcgcggct tccctttcac 720
atgctttgac ttcattatgg cgtaacaaat cgtcggcatt ttgcattcct cgctgcggaa 780
acggaagcgg aagacatccc cgaaagaaac tgcgaactgg gcctttgtgg ccacacgacg 840
aattggtcac gcgccatata gacaattcca tctcagtttt ccacatgttt tgattgctca 900
acgttgtcaa gggactttgt gctgctgtac ttttttcgct tacactttaa tagcttccga 960
agagactcat atccttcata tcctttcttg cagctctgcc atggatctaa agactggctc 1020
tctcctggta gcacccgccg tgaagaactc cagctgctcg catccgcgct actccggcca 1080
caatttcatc gcccacatag ggattgcgga gaccatcgag gcggtgctca tcctggtcct 1140
cacactgggc gtcatcgggg ccaactgcct ggtcatcttc gtcatcaaca accgccgcta 1200
tgcggcctac atacaccaac aggtgggagg ctgagtcact ggcttaacac cttgtacatc 1260
ttatgtctct ccttgtagcc ccgttatctg ctcacctcgc tggccctgaa tgacctgaca 1320
attggtctgc taattacacc attcggtcta atgccggcat tgttccactg ctggccatac 1380
ggagagatct tctgtcagat acaggtgagc gggtcaagag tgggtcaaca ccaggtcaac 1440
aagtggggcg gaagtgcgat aatttgttat ggaacacccg tctcccgttt gcggctgttg 1500
ccactgcggt caattaactc tctcttaaca atgcacagta attactggac ataaatggcg 1560
gtgatggaaa cggacaatcc cgattaacgg gcggaacaat gtgtaatgct agtattaacg 1620
tggtcgagag tgtgtaggat tctgcggtaa gagtcagggt tccgatgcag aaggtgtgac 1680
aagcggaaag ttatcccagc catatggtta ggaagctagt gcacatggaa aaaccattgt 1740
gccattaatt atcatgacgg tggacagggg gcatcgtggt tagggtgtgc aacagaaact 1800
gttagtttat cacttatcaa gtgtcaaaaa gcatgcaacg aaatcaccta gaatccatca 1860
caacaattgt tgcatacatg cttaaaaagg ttgggttttt atgaaaatca gagttgtaaa 1920
cattgtaata aaatccagtt aacactcata caactatttt catgggaacc agactcttgt 1980
tatatatttt atacgatgtc gtttcacaaa atttgaattt tccaactgcc actttagggt 2040
tattggtaaa acatgcacgc attccacgca taactctcac aaatattgaa acacatactg 2100
acgattaata atttattaa aagtttttcg attcgcctcg acaagcgaaa aagtttgccg 2160
aaagcaataa tacaaataat aaacttctaa acgagcacta acgcattttc tggctttttc 2220
gcgttttcca ggccctgctc cgaggggcgc tttcccagca gagcgccgtc atcctggttt 2280
gcatggcagt cgatagatac atgtgtgccc tgcatccgcg ccgctactac cagcactcca 2340
gcaagaaggt aagattccac gtcctatcgc ctcgcatcga tagggaactt ctattttggg 2400
ttgcccggat aaggcacgct cctcggcgat ccttatcctg ttgttcttgc actcgcgctc 2460
gcgctcggaa ggacaataaa caaagctccg gcataatgac gagaaatcac gcgccaataa 2520
aatcgactat aacaatggga agagcaccgc acagcaacta cacaatatgg tcggcgagcc 2580
aattacgcaa gcataaataa ggataatggg ctttaagcag cttgacagtc attaaacagg 2640
caaaaaatgg gtgggcagcg ccaaagagca agtgcaaaac aaagcatcat ggagaatgaa 2700
aaagcaaagt tggcaaaaac caacaaaacc gaaaaaaaaa acaacggggg ccaagtcgaa 2760
```

```
tagacattcg aaaagtgtat aaaatgtgga taaaaaaaag ggagcaaagt tggcctagag 2820
attttttacg cgcaatgact tcggcacagg gcgacgacca cagcaaagtc cttttccagg 2880
acgcaggcag gcaaatgaag cggcacctcc cacggaaacc gcagccagga ccaaagagca 2940
gcaaacgagc tctcaaccaa ccctatttcc cattcccccc ttcgtaattt atggctcccg 3000
gactgtaaac aaacgaaaac taggaggcca acagtcggct aagaacaccg aagtccgaat 3060
gtccttgtcc caaatgaacg ggactgtttc agctgaaatt ggcgatcaat aaatgtatag 3120
tttagaatcc atttcccgca aaattcaagc taacaaacag attttagctt acatatatga 3180
ttatttaacc tttttgattt acagtcttaa tttctcgctg tttgcttggg gattttgcaa 3240
acagttgtaa ttactaaact aactaaaaaa ggaaaatgcc atcaggccaa ataaaacaat 3300
tggaaaacgg gaaagtacca atagtggaaa ataaatttac ataacaattt tgtatgctcc 3360
gaaaatcggc gccatttcgg tttttttcatg ccatctggac gagtgggcga aattggcgca 3420
ctgccacgcc ccatttgtaa aaatgatggg cgcagtgtcg aaaattaata aatggatgct 3480
gaatgcgaaa tactggggat gctgaaaaac caatgctcgg cggaactcgg taatggtaat 3540
tgggaacttt tgactttcgt ttttgcttac tttcaaaagt gccataaatt caaagccagt 3600
cttggacata tctctttatt tatagtgaat ttaacagctt gactcaaaac tcgtcttgtt 3660
gttcgccact cacaatttgt attaaatgca cttgaaaatt caattatttt gttatctagg 3720
atatttgctg actatcctta caacttcagc agtcgccggg ttgtgctcta cgaaaaactc 3780
ctgcaatcct gcaaaaaaaa aaaataatat ttacatttca aaaattgtac ccaactgttt 3840
gagatcgccc tgctttgttt gtttcgttcg ggaaaagttt taaatgtgat agcgtaaatt 3900
aaattaatca cagccatggt ccgtttcaaa gagacacgga caaaagttga gcttcccttt 3960
tttccatgac tttttggcac ttgggatggg gccattttca ttttttcagg tggacaggtg 4020
gacaacattt tcccattcag gatgtcagaa tgctgcattt agtaactcag ccagtcagtg 4080
cgattgtggc agcacttttc cacttccaaa gagcggaagt atcgctttcc ctctgtggtc 4140
tgcacccttt accaagtgat tttggcataa tcaaaagttt gtaattctat ttttggcctc 4200
ttggactccc ttgttgaatg tcagacgaaa tgtgcagtga agcgagtgga attttagcaa 4260
attagaaagt tgaagaggga atttccgact tatctgcacg ccttagtgtg atataagcca 4320
cttccggcta aataaacact taatatgaat tacccaacct tcattcattt tcaattcagt 4380
attggccata ttggcctgaa atgcagggaa tacgtgcctg aagcccttcg ccagtcataa 4440
atcagaccgc acctggcgga aatgtgctcc actaattgtg tattcatttt gcagggctgc 4500
gtggccatac tgagcttaac ctggattatc agcctgacgg tgtttggctt tctggtgctg 4560
cccaaaggtg agctgcccga gattggcccg acatcaatgt gctctcgaac acgtcaaacc 4620
ataaactata tgcgcaggtt actacttcaa caatacgggt ctaatggcct gcgagccctt 4680
ctacagcaag ccatcctaca ggatactgtc cacgtgcgcc ctctactttc ccacgacgat 4740
ggtgctgatg tactgctacg gctccagctt ccacatgagc cgcttccggc taaacgatcc 4800
gaccatgccc ctcacggcgg ccgcccacca cccccatccg cacccacatc ccaccgccgc 4860
ccagcagctg cagatgcacc agcaccagca acatcaccag caggcgggaa tgcactcgca 4920
cctctaccac ggccactcgc atcaccccag ccacccccagt cacccaaacc accccaacca 4980
ccacggccac ccccaccatc atggaccgcc cgttatgggc catctcagca tggccatgag 5040
catgggactg gccggcatgc cgaacatgac aaacaagatt accaaaaagg tgagccagct 5100
gaaccgggtg aactgaggaa gccaggtggc ctgctggcca aaaggatgaa gatggatgaa 5160
ggaggatgaa agaggatgaa gtgggcgggc tccacaacag tttgctcagg cgggcgattt 5220
ggggattaag ttagcagtaa atagaatttt ctcggcatcg ccaaacaaat tgttgggtct 5280
atcagtatgc actcgcagaa cttaaagaag tgttaacagc ttttagtatt tgtgaacggg 5340
ctattcaata ttggtaccag ctaattcaat taaattccag aactcaggta aactaaacta 5400
aaactcactt aaaaaagaac tcgagttgtg caagttactt tacttacttt ccgtgcataa 5460
aaatatttcc agaagccaat attacatttt agcaaaatag aaagcagcag aagttgcatt 5520
tgaaaatata tagaatactc aactatattg ctaataaaca ggaatagaaa tttcgtttgg 5580
gcagtatcta cacataattg aatcaacaat agtcacttta attcaatcgc cgcaaacatt 5640
cccacagatt gttcccatcc aagagaagaa ctccagcgga tccacctcgc gctcgatggc 5700
cgccatttca ctgggcttca tagtgatggt cactccgtgg acgattcagg aaatagtcac 5760
cgcctgcacg ggatccaagg caagtgtcct gtacaaagga ttattgtgtg accgctttgt 5820
gctccgtaca acgcggcgca tacataatga ggtcacgcgt ggctcatcag ttttgatttt 5880
gtgacattta atgttctatg tccgtccaca gctgccgccg ttcctcgatt tcctggtcac 5940
ctggacggcg ctgagcaata gtctgtggaa cccgttcatg tactggctgc tgaactccga 6000
tttccgccgc atgagtcgcc agctgatgcc gaacaaagtg agtatcgggt atccaacagt 6060
gggaagatga atctcctctt ctggcgctaa tttactgcat ctctgatctg gagcgatgtt 6120
cactgtccgt ttcagcagca aatctgccat aagccctgat tgattgattc gcttatccct 6180
attccaaaca ccacgtgcac actttcattg acctccccga acgccaatcc ccaagcccaa 6240
atatgggtat ttcgcccaag tgcatcctgc agctcgcaaa ttttcctcct gcacagcgca 6300
aagtttgcaa aaacaggagc tacatcctcc cgaaagcccc cataaatccg taaccgaagg 6360
```

```
tgtaagctcc agaatgctgg cagactgacc cgggcaaagt aaacaatcaa ccgacgcagc 6420
aacaacaata gcaggtcctg caaagtgatt ttgaattatg aaagagcaaa tctaagacgg 6480
catttacatc ctcgtagccc cacaaacccc atcccatagg tcctgattcc attttggtaa 6540
acgcagtcaa gttgcatgat ttacttccgg agagccgggg tcagtctgtg gaatatttat 6600
ggcccaacct tgtgctctgc cagaataaat gcgcatttaa aatccttatt aatttttgtg 6660
cgagtcggaa tttatgatgc ctccctgccg tttcagtgct tcccccacga ggatacgccc 6720
gagcacaaat caggctgttg tcacatcaat gctaacgatt tcgaaatcac aacgcttccg 6780
attcccccgg agccgccatc atcccgtcct ccgggcggag caggaggagg aggtggaggt 6840
ccatccagtg gaggtcccag ctctggcggc ggagtggcca gttccattgg cggctccgtc 6900
cttggaatct gcggccgctc gcgcacaaac agcctcagcc gcagcgcctc ccagtacatc 6960
cgagggacaa tgggcggagg agtccacaat caccaggcca cggaaacctt ctccaccacg 7020
aggcctgaca ttgagggcct ctccgagaag tactggggcg agatcctgga gagaaccgtc 7080
agctcgggca acctgaatgc catgcagaag cacctgccgc cccacctgcc ctacgcccac 7140
cacctggtcc accagatcca ccagccgcag cagcaccacc aacagcagac cacctccttc 7200
agcaaggcca gcgatctgca gctgaacctg aacctcagcc aagcagccac ggctgccgag 7260
ctgggcaagt tctccaactc ggagccgaag ctgtgcgagc acctctttca cgacgtaaac 7320
tgcgccaaga gcaaggccgc caacggggat ggggagtcct ccggcaagct ggcctgcggt 7380
cgcagcattc cggacatcta gtacgccgag gacgtgatcc tggccaagaa ccaactgctg 7440
gccaggcagg ccaagtgctc gcaccatccg ctgcaccagc aggccaagac gcgtctccat 7500
tcccacacgg gcagcctgcg cttcagccgg attcgcgctg cctgccacag tccgccgccg 7560
gagaagtccg cctttccaat ggccgagtat cgcaagtagg atcacagagg gcactgggat 7620
caaggatgcc gggatacggt actccaggat gtccaggata tggtataacc atgcacagaa 7680
tagggagtga gagcgagcga gtatctataa gattccatag gtgtatgtgc gtgtgagtgt 7740
ctatataata ccgcatgatc aaaccataat atgtgtgttt tctattaggc ttaagatcta 7800
ggaggtcctt tggttgggtt ttcaatttaa agtctctgag attcgcaaaa tacccatggc 7860
ctgtagaaag cgccctgaag cctaactaaa agaatatgaa agatagcacc tatcctcaga 7920
tattccactc tttatataaa gtaatctgtg ttggttataa cctgaatgat tgctgtccat 7980
ctgagaactt tcaataccat tagctagact ttctttacta tacatggatc tggggaatcg 8040
gaattagccg agttaagcag aagttaacca tagtaagtgt tcattcaatg tttgtgatac 8100
cctaagtttc ccagcccaac atacatacct accagtaaca tataatataa tcctagttgt 8160
tgtgcgtgcg tcggctaaat taaacgtagg atatatggaa acagagatat cagactttcc 8220
tatcagaacc cgctcgacat tcggaagccc ttcgccaaat tctccaccac ttagggacta 8280
aggatctacg tagactctga tatcggcgga ccttaaagcg ttgtggttaa tcatagttgt 8340
agttccatag ctcgagtcgt atacaccaaa ttactcactg ttcgacgaat caaaatcaaa 8400
tcctaacttt gtgattgtca agaatttaag tcggttgatg tggacattgc tgttaaatta 8460
atttttattt gtagctaagt aattgactag ttaaagccaa ataaatttac cgtttaagta 8520
agtctaactt tagttaaata aacgagcggt gtgtaacaaa tgtatttcat aggcagagga 8580
ttcaaatcga aataaataa                                              8599
```

<210> 152
<211> 2121
<212> DNA
<213> Drosophila

<400> 152

```
atggatctaa agactggctc tctcctggta gcacccgccg tgaagaactc cagctgctcg 60
catccgcgct actccggcca caatttcatc gcccacatag ggattgcgga gaccatcgag 120
gcggtgctca tcctggtcct cacactgggc gtcatcgggg ccaactgcct ggtcatcttc 180
gtcatcaaca accgccgcta tgcggcctac atacaccaac agccccgtta tctgctcacc 240
tcgctggccc tgaatgacct gacaattggt ctgctaatta caccattcgg tctaatgccg 300
gcattgttcc actgctggcc atacggagag atcttctgtc agatacaggc cctgctccga 360
ggggcgcttt cccagcagag cgccgtcatc ctggtttgca tggcagtcga tagatacatg 420
tgtgccctgc atccgcgccg ctactaccag cactccagca agaagggctg cgtggccata 480
ctgagcttaa cctggattat cagcctgacg gtgtttggct ttctggtgct gcccaaaggt 540
tactacttca acaatacggg tctaatggcc tgcgagccct tctacagcaa gccatcctac 600
aggatactgt ccacgtgcgc cctctacttt cccacgacga tggtgctgat gtactgctac 660
ggctccagct tccacatgag ccgcttccgg ctaaacgatc cgaccatgcc cctcacggcg 720
gccgcccacc acccccatcc gcacccacat cccaccgccg cccagcagct gcagatgcac 780
```

```
cagcaccagc aacatcacca gcaggcggga atgcactcgc acctctacca cggccactcg 840
catcacccca gccaccccag tcacccaaac caccccaacc accacggcca cccccaccat 900
catggaccgc ccgttatggg ccatctcagc atggccatga gcatgggact ggccggcatg 960
ccgaacatga caaacaagat taccaaaaag gtgagccagc tgaaccggat tgttcccatc 1020
caagagaaga actccagcgg atccacctcg cgctcgatgg ccgccatttc actgggcttc 1080
atagtgatgg tcactccgtg gacgattcag gaaatagtca ccgcctgcac gggatccaag 1140
gcaagtgtcc tgtacaaagg attattgtgt gaccgctttc tgccgccgtt cctcgatttc 1200
ctggtcacct ggacggcgct gagcaatagt ctgtggaacc cgttcatgta ctggctgctg 1260
aactccgatt tccgccgcat gagtcgccag ctgatgccga acaaatgctt cccccacgag 1320
gatacgcccg agcacaaatc aggctgttgt cacatcaatg ctaacgattt cgaaatcaca 1380
acgcttccga ttcccccgga gccgccatca tcccgtcctc cgggcggagc aggaggagga 1440
ggtggaggtc catccagtgg aggtcccagc tctggcggcg gagtggccag ttccattggc 1500
ggctccgtcc ttggaatctg cggccgctcg cgcacaaaca gcctcagccg cagcgcctcc 1560
cagtacatcc gagggacaat gggcggagga gtccacaatc accaggccac ggaaaccttc 1620
tccaccacga ggcctgacat tgagggcctc tccgagaagt actggggcga gatcctggag 1680
agaaccgtca gctcgggcaa cctgaatgcc atgcagaagc acctgccgcc ccacctgccc 1740
tacgcccacc acctggtcca ccagatccac cagccgcagc agcaccacca acagcagacc 1800
acctccttca gcaaggccag cgatctgcag ctgaacctga acctcagcca agcagccacg 1860
gctgccgagc tgggcaagtt ctccaactcg gagccgaagc tgtgcgagca cctctttcac 1920
gactacgccg aggacgtgat cctggccaag aaccaactgc tggccaggca ggccaagtgc 1980
tcgcaccatc cgctgcacca gcaggccaag acgcgtctcc attcccacac gggcagcctg 2040
cgcttcagcc ggattcgcgc tgcctgccac agtccgccgc cggagaagtc cgcctttcca 2100
atggccgagt atcgcaagta g                                            2121
```

<210> 153
<211> 706
<212> PRT
<213> Drosophila

<400> 153

```
Met Asp Leu Lys Thr Gly Ser Leu Leu Val Ala Pro Ala Val Lys Asn
 1               5                  10                  15
Ser Ser Cys Ser His Pro Arg Tyr Ser Gly His Asn Phe Ile Ala His
            20                  25                  30
Ile Gly Ile Ala Glu Thr Ile Glu Ala Val Leu Ile Leu Val Leu Thr
        35                  40                  45
Leu Gly Val Ile Gly Ala Asn Cys Leu Val Ile Phe Val Ile Asn Asn
    50                  55                  60
Arg Arg Tyr Ala Ala Tyr Ile His Gln Gln Pro Arg Tyr Leu Leu Thr
65                  70                  75                  80
Ser Leu Ala Leu Asn Asp Leu Thr Ile Gly Leu Leu Ile Thr Pro Phe
                85                  90                  95
Gly Leu Met Pro Ala Leu Phe His Cys Trp Pro Tyr Gly Glu Ile Phe
            100                 105                 110
Cys Gln Ile Gln Ala Leu Leu Arg Gly Ala Leu Ser Gln Gln Ser Ala
        115                 120                 125
Val Ile Leu Val Cys Met Ala Val Asp Arg Tyr Met Cys Ala Leu His
    130                 135                 140
Pro Arg Arg Tyr Tyr Gln His Ser Ser Lys Lys Gly Cys Val Ala Ile
145                 150                 155                 160
Leu Ser Leu Thr Trp Ile Ile Ser Leu Thr Val Phe Gly Phe Leu Val
                165                 170                 175
Leu Pro Lys Gly Tyr Tyr Phe Asn Asn Thr Gly Leu Met Ala Cys Glu
            180                 185                 190
Pro Phe Tyr Ser Lys Pro Ser Tyr Arg Ile Leu Ser Thr Cys Ala Leu
        195                 200                 205
Tyr Phe Pro Thr Thr Met Val Leu Met Tyr Cys Tyr Gly Ser Ser Phe
    210                 215                 220
His Met Ser Arg Phe Arg Leu Asn Asp Pro Thr Met Pro Leu Thr Ala
225                 230                 235                 240
```

```
Ala Ala His His Pro His Pro His Pro His Pro Thr Ala Ala Gln Gln
                245                 250                 255
Leu Gln Met His Gln His Gln Gln His His Gln Gln Ala Gly Met His
            260                 265                 270
Ser His Leu Tyr His Gly His Ser His His Pro Ser His Pro Ser His
        275                 280                 285
Pro Asn His Pro Asn His His Gly His Pro His His His Gly Pro Pro
    290                 295                 300
Val Met Gly His Leu Ser Met Ala Met Ser Met Gly Leu Ala Gly Met
305                 310                 315                 320
Pro Asn Met Thr Asn Lys Ile Thr Lys Lys Val Ser Gln Leu Asn Arg
                325                 330                 335
Ile Val Pro Ile Gln Glu Lys Asn Ser Ser Gly Ser Thr Ser Arg Ser
            340                 345                 350
Met Ala Ala Ile Ser Leu Gly Phe Ile Val Met Val Thr Pro Trp Thr
        355                 360                 365
Ile Gln Glu Ile Val Thr Ala Cys Thr Gly Ser Lys Ala Ser Val Leu
    370                 375                 380
Tyr Lys Gly Leu Leu Cys Asp Arg Phe Leu Pro Pro Phe Leu Asp Phe
385                 390                 395                 400
Leu Val Thr Trp Thr Ala Leu Ser Asn Ser Leu Trp Asn Pro Phe Met
                405                 410                 415
Tyr Trp Leu Leu Asn Ser Asp Phe Arg Arg Met Ser Arg Gln Leu Met
            420                 425                 430
Pro Asn Lys Cys Phe Pro His Glu Asp Thr Pro Glu His Lys Ser Gly
        435                 440                 445
Cys Cys His Ile Asn Ala Asn Asp Phe Glu Ile Thr Thr Leu Pro Ile
    450                 455                 460
Pro Pro Glu Pro Pro Ser Ser Arg Pro Pro Gly Gly Ala Gly Gly Gly
465                 470                 475                 480
Gly Gly Gly Pro Ser Ser Gly Gly Pro Ser Ser Gly Gly Gly Val Ala
                485                 490                 495
Ser Ser Ile Gly Gly Ser Val Leu Gly Ile Cys Gly Arg Ser Arg Thr
            500                 505                 510
Asn Ser Leu Ser Arg Ser Ala Ser Gln Tyr Ile Arg Gly Thr Met Gly
        515                 520                 525
Gly Gly Val His Asn His Gln Ala Thr Glu Thr Phe Ser Thr Thr Arg
    530                 535                 540
Pro Asp Ile Glu Gly Leu Ser Glu Lys Tyr Trp Gly Glu Ile Leu Glu
545                 550                 555                 560
Arg Thr Val Ser Ser Gly Asn Leu Asn Ala Met Gln Lys His Leu Pro
                565                 570                 575
Pro His Leu Pro Tyr Ala His His Leu Val His Gln Ile His Gln Pro
            580                 585                 590
Gln Gln His His Gln Gln Gln Thr Thr Ser Phe Ser Lys Ala Ser Asp
        595                 600                 605
Leu Gln Leu Asn Leu Asn Leu Ser Gln Ala Ala Thr Ala Ala Glu Leu
    610                 615                 620
Gly Lys Phe Ser Asn Ser Glu Pro Lys Leu Cys Glu His Leu Phe His
625                 630                 635                 640
Asp Tyr Ala Glu Asp Val Ile Leu Ala Lys Asn Gln Leu Leu Ala Arg
                645                 650                 655
Gln Ala Lys Cys Ser His His Pro Leu His Gln Gln Ala Lys Thr Arg
            660                 665                 670
Leu His Ser His Thr Gly Ser Leu Arg Phe Ser Arg Ile Arg Ala Ala
        675                 680                 685
Cys His Ser Pro Pro Pro Glu Lys Ser Ala Phe Pro Met Ala Glu Tyr
    690                 695                 700
Arg Lys
705
```

<210> 154
<211> 3332
<212> DNA
<213> Drosophila

<400> 154
gataaagccc aattacagaa cctaatacat acacctaaaa tgactaattg acatccattg 60
tgcgtattac ataataatga ataatttaag tgatagttgt ttgaacttac agaacacttc 120
ttagtattgc ggactcatcc catccaactg ttgactcatt tcaaattcta tttaattata 180
tttctggtcc tggcacgaac aataaattga gtttatatgt attacgcccc ttcttgtggt 240
ctcactttat gaaaaagtta attgcgtgta gttgttggct ttgctagatg gatttttatt 300
gtgccgcttt actgtctgct tttttcttca ttttccactc gcctttcatt tttcatccgc 360
cggaaagcga atcatttttt cgcacgctga gcagtgcaaa gaaagcaaac acaagcaaac 420
gagcagaatt tcatatagtt gattgcattg tctgcaggaa gatgttcgca cgatgatgca 480
ggacatgcag gactaaggac ttcgaagcga atggtgtggg ctggggcagg gatagtctca 540
tgaaatgctg atagcatgcg ctcaatattg tcagtaaata atacgcgctt gttagagcca 600
actgcatacc aactaaactg aacacagtga gaataaacat aacagaaagt tgaatttgct 660
taatatataa tttcataaat ttgtcacacc aaaattcgga tttaaattaa aagtataatt 720
atattaataa cattatttaa tataattata attatatata aaaatatatt tttatatgta 780
tgtttcattt ggttgctgaa aaccagtaac taactaatct ttgtgggcta tccaaatctc 840
ttcattgcag aaaccgcagc gtgcaacatg tgacgtcaca ggcataaagt atggagagca 900
tattaaaata tttgccattt agctggttta gcatgcattg ttgcattgac tgccacataa 960
caagtggaaa ttagcaggca gctactgtag acgaatcgaa atggaaggtg gatggtggcg 1020
aggaggagga ggaggaggac gtttgggcgg gaaagccatc atggaaggac actcgacacc 1080
aaatggggca gctgcaagcc acagaaacaa cagcacaagg accaatatcg cgaccaacgg 1140
ctgtgcccat tcgggcatcc tgttatttgt gctgacagct atgacactaa cgagcttaat 1200
aacgcccaca gagcagctgg cagtggcgcc aaatggcacc acactgcatc aactggagtc 1260
cgtggagtcc gagtcgtatc catccataaa tggcactcag aatgaaacaa tggtgaccag 1320
tgtgcgaccc cacctggacc acaggaatcg accgacgcag cagaatggca gtcactattt 1380
ggagtacgac gatgatggcc cggactgttc gtacagctac aacttcatcc tgaagctcat 1440
cacgatgatc ctgtacgcac tggtctgcat cattggactc tttggcaaca ccctggtgat 1500
ctatgtggtg atgaggttct ccaagatgca gaccgtaacc aacatataca tactgaatct 1560
ggctatcgcg gatgagtgct tcctgatcgg cattcccttc ctgctctaca caatgcaggt 1620
gggcaactgg cccttcggca actatatgtg caaggcctac atggtgagca cctcgatcac 1680
ctccttcacc tcctcgatct tcctgttgat catgtcggcg gatcgctaca tagccgtttg 1740
ccatcccata tcctcgcctc gctaccgaac gccctttgta tccaagttgg tttcggcctt 1800
cgcctggatg acatccgtgc tgctgatgct gccggttatc ctttttgcca gcaccgtgca 1860
gtcgagcaac ggcaatgtgt cctgcaacat cgagtggcca gacactcaga actcgcacac 1920
cgactccacc ttcattttgt actcgctggt cttgggattc gccactccac tgacttttat 1980
cctggtgttc tactgcctgg tgatcaggaa acttcacacc gtgggaccga agcacaagtc 2040
taaggagaag aagcgctctc acaggaaggt caccaagttg gtgctcacgg taataatata 2100
tgacatttcg attagttgtg gaactaacta ctacctgacc actcttgtgt tcagtttgat 2160
aaatattatt attacattat aaatatatat ttttttttga cattttgaat tcgactataa 2220
attaattgtt ttaaattgat atatagaaag aattttataa tatttattca atacaattta 2280
tggttttttg atttctccac aggtcataag tgcgtacata ttttgttggc ttccacactg 2340
gatttcacag gtaactaatt gtatacatta cagtctaaaa acatttatct ttgtgtactt 2400
taaaaggttt taaaaccata agctatttgc ctactattat ccttttggat tgatttttaa 2460
attattatag ctaaaacaaa ttacgagatt tgtggaaatt gacatgaact ggacttcttt 2520
ggcagatcat ttgtagattc gctttttatt gactattttc caggtggctt tgatcagctc 2580
cgctcctcaa cgctgtgcat ctcgtctgga gctggccgtc ttcctggcct gcggatgcct 2640
cagctactcc aactcggcca tgaacccaat actgtacgcc tttttgagcg ataacttcaa 2700
gaagagcttc atgaaggcct gcacgtgtgc tgcccgcaag gatgtgaatg cccagttgca 2760
gctggagaac agtttcttcc ccaagtttgg caagggcagg caatcggagc gtcttcttgg 2820
tggcaatgga aaaggtggcg cccagcgtgg ggcattaacc aagaagaagt gcttggcgac 2880
gagaaacaac aatgctccga tggccactac aacgacgacg acaaccacca caacgggcac 2940
ggatgcagtg acctgtctcc agccgccagt acaccaggtg ccagccgaga tccaggtggg 3000
aaatccggcc accgtgctgg tggtcaatgc tgagaccaac aactgtaaac cgcccgtgct 3060
ccacacggac ttataagacc gggctccctc gatgcccctg gaaacggtgg tcttcatcgc 3120

gcggaggtga gaccatcagg ttgagctgga cctcgatacg gccatcgatt gccaggcaat 3180
agcgaggcag ccggaatgct tgctataggt cgcctaaaat gatattaaca ttttatacca 3240
atatggcaga gctttgcaaa aacgggtaaa cttaatccca tttttatatg atcgaatctc 3300
agcaagaaga aatttttaaa aacccgaata aa 3332

<210> 155
<211> 1332
<212> DNA
<213> Drosophila

<400> 155
atggaaggtg gatggtggcg aggaggagga ggaggaggac gtttgggcgg gaaagccatc 60
atggaaggac actcgacacc aaatggggca gctgcaagcc acagaaacaa cagcacaagg 120
accaatatcg cgaccaacgg ctgtgcccat tcgggcatcc tgttatttgt gctgacagct 180
atgacactaa cgagcttaat aacgcccaca gagcagctgg cagtggcgcc aaatggcacc 240
acactgcatc aactggagtc cgtggagtcc gagtcgtatc catccataaa tggcactcag 300
aatgaaacaa tggtgaccag tgtgcgaccc cacctggacc acaggaatcg accgacgcag 360
cagaatggca gtcactattt ggagtacgac gatgatggcc cggactgttc gtacagctac 420
aacttcatcc tgaagctcat cacgatgatc ctgtacgcac tggtctgcat cattggactc 480
tttggcaaca ccctggtgat ctatgtggtg atgaggttct ccaagatgca gaccgtaacc 540
aacatataca tactgaatct ggctatcgcg gatgagtgct tcctgatcgg cattcccttc 600
ctgctctaca caatgcaggt gggcaactgg cccttcggca actatatgtg caaggcctac 660
atggtgagca cctcgatcac ctccttcacc tcctcgatct tcctgttgat catgtcggcg 720
gatcgctaca tagccgtttg ccatcccata tcctcgcctc gctaccgaac gccctttgta 780
tccaagttgg tttcggcctt cgcctggatg acatccgtgc tgctgatgct gccggttatc 840
ctttttgcca gcaccgtgca gtcgagcaac ggcaatgtgt cctgcaacat cgagtggcca 900
gacactcaga actcgcacac cgactccacc ttcattttgt actcgctggt cttgggattc 960
gccactccac tgacttttat cctggtgttc tactgcctgg tgatcaggaa acttcacacc 1020
gtgggaccga agcacaagtc taaggagaag aagcgctctc acaggaaggt caccaagttg 1080
gtgctcacgg taataatata tgacatttcg attagttgtg gaactaacta ctacctgacc 1140
actcttgtgt tcagtttgat aaatattatt attacattat aaatatatat tttttttga 1200
cattttgaat tcgactataa attaattgtt ttaaattgat atatagaaag aattttataa 1260
tatttattca atacaattta tggtttttg atttctccac aggtcataag tgcgtacata 1320
ttttgttggc tt 1332

<210> 156
<211> 393
<212> PRT
<213> Drosophila

<400> 156
Met Glu Gly Gly Trp Trp Arg Gly Gly Gly Gly Gly Gly Arg Leu Gly
1 5 10 15
Gly Lys Ala Ile Met Glu Gly His Ser Thr Pro Asn Gly Ala Ala Ala
20 25 30
Ser His Arg Asn Asn Ser Thr Arg Thr Asn Ile Ala Thr Asn Gly Cys
35 40 45
Ala His Ser Gly Ile Leu Leu Phe Val Leu Thr Ala Met Thr Leu Thr
50 55 60
Ser Leu Ile Thr Pro Thr Glu Gln Leu Ala Val Ala Pro Asn Gly Thr
65 70 75 80
Thr Leu His Gln Leu Glu Ser Val Glu Ser Glu Ser Tyr Pro Ser Ile
85 90 95
Asn Gly Thr Gln Asn Glu Thr Met Val Thr Ser Val Arg Pro His Leu
100 105 110
Asp His Arg Asn Arg Pro Thr Gln Gln Asn Gly Ser His Tyr Leu Glu
115 120 125
Tyr Asp Asp Asp Gly Pro Asp Cys Ser Tyr Ser Tyr Asn Phe Ile Leu
130 135 140
Lys Leu Ile Thr Met Ile Leu Tyr Ala Leu Val Cys Ile Ile Gly Leu

145 150 155 160
Phe Gly Asn Thr Leu Val Ile Tyr Val Val Met Arg Phe Ser Lys Met
165 170 175
Gln Thr Val Thr Asn Ile Tyr Ile Leu Asn Leu Ala Ile Ala Asp Glu
180 185 190
Cys Phe Leu Ile Gly Ile Pro Phe Leu Leu Tyr Thr Met Gln Val Gly
195 200 205
Asn Trp Pro Phe Gly Asn Tyr Met Cys Lys Ala Tyr Met Val Ser Thr
210 215 220
Ser Ile Thr Ser Phe Thr Ser Ser Ile Phe Leu Leu Ile Met Ser Ala
225 230 235 240
Asp Arg Tyr Ile Ala Val Cys His Pro Ile Ser Ser Pro Arg Tyr Arg
245 250 255
Thr Pro Phe Val Ser Lys Leu Val Ser Ala Phe Ala Trp Met Thr Ser
260 265 270
Val Leu Leu Met Leu Pro Val Ile Leu Phe Ala Ser Thr Val Gln Ser
275 280 285
Ser Asn Gly Asn Val Ser Cys Asn Ile Glu Trp Pro Asp Thr Gln Asn
290 295 300
Ser His Thr Asp Ser Thr Phe Ile Leu Tyr Ser Leu Val Leu Gly Phe
305 310 315 320
Ala Thr Pro Leu Thr Phe Ile Leu Val Phe Tyr Cys Leu Val Ile Arg
325 330 335
Lys Leu His Thr Val Gly Pro Lys His Lys Ser Lys Glu Lys Lys Arg
340 345 350
Ser His Arg Lys Val Thr Lys Leu Val Leu Thr Val Ile Ile Tyr Asp
355 360 365
Ile Ser Ile Ser Cys Gly Thr Asn Tyr Tyr Leu Thr Thr Leu Val Phe
370 375 380
Ser Leu Ile Asn Ile Ile Ile Thr Leu
385 390

<210> 157
<211> 5453
<212> DNA
<213> Drosophila

<400> 157
gataaaagca ttacattttc gagtctgcac gatcgtatca tcagttgttt gcaggacctt 60
gagttcttca tacaaagttt tagctgctca tttttaatct gttccttgtg cagagattac 120
acttctccga aaagagccga cacgtaaatt ggaaaagggt tagtgcaact gttgttgcaa 180
ttacaattca tcacatagcc cgtagttcct cttcctgctg cacagaaaga aattggtatg 240
tggcgatcgt tgcaatgcca tgcaattcaa tgcccaattc aaaaatattt gaaatgtgta 300
aacaactcac ttgttttctt ttctgcgtgc atgtggtgct cctggccatt tgccatttcg 360
ccaggcataa tcccgacaat cggtgcgatg tgggtgatac gagtgcgaat gcgaatgcga 420
atgccaatga tatgtgtatt gtatgtatgt gcatacatat ggcgtcgtag cgtgtgggcg 480
tgatgtttgc actccggtgc agcgggacta taaataagcc gggtatttat taggggcaat 540
cataaactgc gttaacgacg ggcactctgc caacgagaat aagatccggt aaacaaaggc 600
accggtcgac gtacgtctca ttgtgcgata agctttcaat taaatctgac ccggataacg 660
gggcgcatgg gtgcgggaaa tcggaaatcg gaaacgaaaa cgaaaacgga agcggaaatc 720
gaaatcgaaa tggaaagaga tcaattttca attgcagcca atgcatgcat gtccatgggt 780
ccaatgctaa tatcgaaaga caaagctccg tgttctggag gtcgtgttcg ccatgcagac 840
tcattggttg gttgggttta actcgtacaa agtttgatgg gtttgggatt gttttttttt 900
ttttttaga acgatataac tcgatataat gctggtgcgt aaataacaac gatgttcatt 960
gttcgctttg cagcacattt actatgcggt cgatggcaaa atgaccctcc tgtcgaacat 1020
tctcgactgc ggaggctgta tttccgccca gcgcttcacc cgcctgctgc gccagtccgg 1080
ctcatcagga ccatccccat ctgcaccgac ggccggaaca tttgaatcaa aatccatgct 1140
ggagccaaca tcctcgcaca gcctggcgac cggacgcgtg ccactactgc acgatttcga 1200
tgcctcgaca acggaatcgc cgggaacgta tgtcctcgac ggtgtcgcca gggtggccca 1260

attggccctg gagcccaccg tcatggacgc actgcccgat tcggacacgg aacaggttct 1320
cggtaatctc aacagcagcg cgccctggaa tctaacgctt gcgtcggcag cggccaccaa 1380
tttcgagaac tgctcggcgc tttttgtcaa ttacacccta ccacaaacag gtacctaaag 1440
cctaataact tttaaatcat attgcagtat tttttttttt ttttttgttg ccgggatagg 1500
actctactgc aattggacct gggacacatt gctctgctgg ccacccactc cggctggagt 1560
ccttgcacgg atgaattgtc ctggcggctt tcatggcgta gatacgcgca gtaagtaact 1620
gaaatacact ctaaaaaaaa agcagccgaa aaagaacatg ctgagtacaa ttagcataat 1680
taatcgttga agctgtttag ctgtttagca gccacaagaa gctttatgac ttctgccggc 1740
ttattaggag tcaaaagaca tgtcctacgg tgtttatccc atccacacgg catatctggc 1800
ggtcacaatc atacggatca ttagctggga aacggttcaa gtggtgcacg gcaaacaatt 1860
acttgattcc gtattcatcg aaaggatgtc aatatttaca tgcatgactg gggcgtcca 1920
atcaaccgca ctcccagtcg gggccgcccc tccttggctt ccatcgccgg gggctgtcca 1980
ggtgtttcat gtgcaccctc gagctccgcc agttgaacag ggtttaaaac accagccaat 2040
cgattgctta attacgatcc taaatgcttg ctctacccct ccgataaatt aaacgttctg 2100
caatggggaa ttgctattta ataaccatgg gcgatataca ctgagcgaaa agattctaag 2160
tagctgactt accacccaat attttacaaa cttaaaatgc tcttccatta tcactagtga 2220
atgtattatc tgcgcttgtg gtgcaaaatc taggcatatc gaactatatc tgttaagaaa 2280
taattattcg agtgcattgt tgagctacgg cctcatttct ccgggccgtg tctagatttt 2340
tagccccaat attttcaact ggacgaggtc ggtcgcccat tcgattggca taattgcact 2400
tgtcagtcag tgtgtcaaag ctcgattact gtcattaatc caacccaacc tccctctttt 2460
tcggttcgca gaattcgcca tccgaaagtg tgagctggat ggtcgatggg gcagcaggcc 2520
aaatgccacg gaggtgaatc cgccgggatg gacggactac gggccgtgtt acaagccgga 2580
gattatccgt ctcatgcagc agatgggcag caaggacttc gatgtgagtt gaattctttc 2640
cacactacat cataatacct gtttcagtta tagttaagag cattgcttaa atgccaacca 2700
aaatcccaaa tcaagggtat ggttttgaaa ttacctaata gggttccaag ttttttagac 2760
accctaaaaa gtgatttcaa aggcctgatt acatttgttg atagtggttt atagtagtag 2820
ttttaattcc atttcttcca actggtttgc gatatccctt aggcctacat agacattgcc 2880
aggaggactc gaaccctgga gatcgtgggc ctctgcctct ccctgttcgc ccttatagtt 2940
tccctgctga tcttctgcac atttcgctcg ctgcgaaaca atcgcaccaa gatccacaag 3000
aatcttttcg tcgccatggt gctgcaggtg atcattcgcc tgaccttgta tctcgaccaa 3060
ttccggcggg gaaacaagga ggcggccacc aacacgagtc tctctgtcat tgagaacacg 3120
gtaggcttaa tattcccgca taaggataag gtataaggct tatgttcgat attatatgtt 3180
atgttctaca gccctatttg tgcgaagcat cctatgtact tctggagtac gctcgtaccg 3240
ccatgttcat gtggatgttc atcgagggcc tttacctgca caacatggtc accgtggccg 3300
ttttccaggg cagctttccc ctcaagttct tctcgcgact cggctggtgt gtgcccattc 3360
tgatgaccac cgtgtgggcg agatgcacgg tcatgtatat ggacacctcg ctgggcgaat 3420
gcttgtggaa ctataatctc acgccctact actggatcct cgaggggcca cgactagcgg 3480
tcatactggt gggtaatatc aactgttgat cgctttaatc acctattcac ccaactgaaa 3540
acatccgcat gcagctaaac ttctgtttcc tggtgaacat tatccgagtg ctggtaatga 3600
agctgcgtca atcgcaggcc agcgatatag aacagactcg caaggcagtt agagcggcta 3660
tagtcctact accacttttg ggtataacca atctcctgca ccagctggct cctctgaaaa 3720
cggccacgaa cttcgcggtc tggtcgtatg gcacccactt tctcacctcg tttcagggat 3780
tttttatagc gctaatttac tgctttctaa atggcgaggt aagacatcaa gtcactcagt 3840
tgtccatttg cctaaaaccc tcactatttc caaccaaaat atatattttt taggttcgtg 3900
ccgtgctact aaagagtctg gccacccagc tgtcggtgcg aggtcatccg gaatgggcgc 3960
cgaaaagggc atctatgtac tcgggtgctt ataacacggc gccggatacg gatgcagtgc 4020
agcctgcagg agatccatcg gccactggaa agcggtaagc atccgctttg tgcacacccc 4080
accttgaaac tccttacatc aatttaaaac tcccagaata tcaccgccga ataaaaggct 4140
gaatggaaga aagccgagca gtgccagcat tgtgatgatt cacgagcctc aacagcgcca 4200
gcgactgatg ccccggctgc aaaacaaggc gcgggaaaag ggcaaggacc gggtggagaa 4260
gacggatgcg gaagcggagc cggatccgac catctcccac attcacagca aggaggcggg 4320
cagcgcgaga tcgcgaactc gcggctccaa gtggataatg ggcatctgct tccggggtca 4380
aaaggtacta agagtaccgt cagcgtcatc cgtgccaccc gagtcagttg tatttgagtt 4440
gtcagagcag tagaccaaaa cctagcctaa ctaatagatg tagccatagc cgtagccaaa 4500
tagccgtaga agtagccaac gccatgtagc ccaccaccac agtagccagc acccacacaa 4560
tcggaccgga aatgatccca aagataaccc agaacccgac tattcctgca tctcccatcg 4620
acccaccaag caagccagct ggaaacagca catatgtaaa ataccccaac aaaatatcgc 4680
agttcttata cccaagatca ttttgaacag aaagtaagct cattgttttg taattcaaat 4740
taaaccaata gacatttaat ttaaatagtt attttcgctt gccttcggtt cttgctccac 4800
tgtacttatg taaatacata catacctta tttgtgtaag ccatatccaa accctttcca 4860

aaatctagaa atctcctatc ctttcgaaat cggctaatcc tacgaaaatt caaatttccc 4920
tacaagtttt acaaaatatt tttgatgaca ataaatttcc aatagcgtca ccttctgttc 4980
tggacacgca aatacattca aaaacactcg cacacggttc ttactcgatg aatttttcag 5040
caaaacttca gcggcatgtt cttttacttc tactattcat atttaaatat tagattgaaa 5100
tatttaagtg aattaacaca aacttgcggt aactgcgcag tgtggccagg ctttgtgagc 5160
cattagaact tattacataa cataattgaa aagtgacatt gtttaaagga actcaccttt 5220
aatcaactgt tgaacttgct gttcatcgga ttccgttatt aagttttaag atttgggttt 5280
acttgtattt acattttctt taccactttg attagcttac actttccgtt cctgagctgg 5340
tcacactgcc agctacaatc gatcacgttc actgcaactg ccttgaattt tgctagctgc 5400
gctttgttgt gcaacaaaaa ctttgccaca atcgatattt aactaatggg gtc 5453

<210> 158
<211> 1902
<212> DNA
<213> Drosophila

<400> 158
atgaccctcc tgtcgaacat tctcgactgc ggaggctgta tttccgccca gcgcttcacc 60
cgcctgctgc gccagtccgg ctcatcagga ccatccccat ctgcaccgac ggccggaaca 120
tttgaatcaa aatccatgct ggagccaaca tcctcgcaca gcctggcgac cggacgcgtg 180
ccactactgc acgatttcga tgcctcgaca acggaatcgc cgggaacgta tgtcctcgac 240
ggtgtcgcca gggtggccca attggccctg gagcccaccg tcatggacgc actgcccgat 300
tcggacacgg aacaggttct cggactctac tgcaattgga cctgggacac attgctctgc 360
tggccaccca ctccggctgg agtccttgca cggatgaatt gtcctggcgg ctttcatggc 420
gtagatacgc gcaaattcgc catccgaaag tgtgagctgg atggtcgatg ggcagcaggg 480
ccaaatgcca cggaggtgaa tccgccggga tggacggact acgggccgtg ttacaagccg 540
gagattatcc gtctcatgca gcagatgggc agcaaggact tcgatgccta catagacatt 600
gccaggagga ctcgaaccct ggagatcgtg ggcctctgcc tctccctgtt cgcccttata 660
gtttccctgc tgatcttctg cacatttcgc tcgctgcgaa acaatcgcac caagatccac 720
aagaatcttt tcgtcgccat ggtgctgcag gtgatcattc gcctgacctt gtatctcgac 780
caattccggc ggggaaacaa ggaggcggcc accaacacga gtctctctgt cattgagaac 840
acgccctatt tgtgcgaagc atcctatgta cttctggagt acgctcgtac cgccatgttc 900
atgtggatgt tcatcgaggg cctttacctg cacaacatgg tcaccgtggc cgttttccag 960
ggcagctttc ccctcaagtt cttctcgcga ctcggctggt gtgtgcccat tctgatgacc 1020
accgtgtggg cgagatgcac ggtcatgtat atggacacct cgctgggcga atgcttgtgg 1080
aactataatc tcacgcccta ctactggatc ctcgaggggc cacgactagc ggtcatactg 1140
ctaaacttct gtttcctggt gaacattatc cgagtgctgg taatgaagct gcgtcaatcg 1200
caggccagcg atatagaaca gactcgcaag gcagttagag cggctatagt cctactacca 1260
cttttgggta taaccaatct cctgcaccag ctggctcctc tgaaaacggc cacgaacttc 1320
gcggtctggt cgtatggcac ccactttctc acctcgtttc agggattttt tatagcgcta 1380
atttactgct ttctaaatgg cgaggttcgt gccgtgctac taaagagtct ggccacccag 1440
ctgtcggtgc gaggtcatcc ggaatgggcg ccgaaaaggg catctatgta ctcgggtgct 1500
tataacacgg cgccggatac ggatgcagtg cagcctgcag gagatccatc ggccactgga 1560
aagcgaatat caccgccgaa taaaaggctg aatggaagaa agccgagcag tgccagcatt 1620
gtgatgattc acgagcctca acagcgccag cgactgatgc cccggctgca aaacaaggcg 1680
cgggaaaagg gcaaggaccg ggtggagaag acggatgcgg aagcggagcc ggatccgacc 1740
atctcccaca ttcacagcaa ggaggcgggc agcgcgagat cgcgaactcg cggctccaag 1800
tggataatgg gcatctgctt ccggggtcaa aaggtactaa gagtaccgtc agcgtcatcc 1860
gtgccacccg agtcagttgt atttgagttg tcagagcagt ag 1902

<210> 159
<211> 633
<212> PRT
<213> Drosophila

<400> 159
Met Thr Leu Leu Ser Asn Ile Leu Asp Cys Gly Gly Cys Ile Ser Ala
1   5   10   15

Gln Arg Phe Thr Arg Leu Leu Arg Gln Ser Gly Ser Ser Gly Pro Ser
20 25 30
Pro Ser Ala Pro Thr Ala Gly Thr Phe Glu Ser Lys Ser Met Leu Glu
35 40 45
Pro Thr Ser Ser His Ser Leu Ala Thr Gly Arg Val Pro Leu Leu His
50 55 60
Asp Phe Asp Ala Ser Thr Thr Glu Ser Pro Gly Thr Tyr Val Leu Asp
65 70 75 80
Gly Val Ala Arg Val Ala Gln Leu Ala Leu Glu Pro Thr Val Met Asp
85 90 95
Ala Leu Pro Asp Ser Asp Thr Glu Gln Val Leu Gly Leu Tyr Cys Asn
100 105 110
Trp Thr Trp Asp Thr Leu Leu Cys Trp Pro Pro Thr Pro Ala Gly Val
115 120 125
Leu Ala Arg Met Asn Cys Pro Gly Gly Phe His Gly Val Asp Thr Arg
130 135 140
Lys Phe Ala Ile Arg Lys Cys Glu Leu Asp Gly Arg Trp Gly Ser Arg
145 150 155 160
Pro Asn Ala Thr Glu Val Asn Pro Pro Gly Trp Thr Asp Tyr Gly Pro
165 170 175
Cys Tyr Lys Pro Glu Ile Ile Arg Leu Met Gln Gln Met Gly Ser Lys
180 185 190
Asp Phe Asp Ala Tyr Ile Asp Ile Ala Arg Arg Thr Arg Thr Leu Glu
195 200 205
Ile Val Gly Leu Cys Leu Ser Leu Phe Ala Leu Ile Val Ser Leu Leu
210 215 220
Ile Phe Cys Thr Phe Arg Ser Leu Arg Asn Asn Arg Thr Lys Ile His
225 230 235 240
Lys Asn Leu Phe Val Ala Met Val Leu Gln Val Ile Ile Arg Leu Thr
245 250 255
Leu Tyr Leu Asp Gln Phe Arg Arg Gly Asn Lys Glu Ala Ala Thr Asn
260 265 270
Thr Ser Leu Ser Val Ile Glu Asn Thr Pro Tyr Leu Cys Glu Ala Ser
275 280 285
Tyr Val Leu Leu Glu Tyr Ala Arg Thr Ala Met Phe Met Trp Met Phe
290 295 300
Ile Glu Gly Leu Tyr Leu His Asn Met Val Thr Val Ala Val Phe Gln
305 310 315 320
Gly Ser Phe Pro Leu Lys Phe Phe Ser Arg Leu Gly Trp Cys Val Pro
325 330 335
Ile Leu Met Thr Thr Val Trp Ala Arg Cys Thr Val Met Tyr Met Asp
340 345 350
Thr Ser Leu Gly Glu Cys Leu Trp Asn Tyr Asn Leu Thr Pro Tyr Tyr
355 360 365
Trp Ile Leu Glu Gly Pro Arg Leu Ala Val Ile Leu Leu Asn Phe Cys
370 375 380
Phe Leu Val Asn Ile Ile Arg Val Leu Val Met Lys Leu Arg Gln Ser
385 390 395 400
Gln Ala Ser Asp Ile Glu Gln Thr Arg Lys Ala Val Arg Ala Ala Ile
405 410 415
Val Leu Leu Pro Leu Leu Gly Ile Thr Asn Leu Leu His Gln Leu Ala
420 425 430
Pro Leu Lys Thr Ala Thr Asn Phe Ala Val Trp Ser Tyr Gly Thr His
435 440 445
Phe Leu Thr Ser Phe Gln Gly Phe Phe Ile Ala Leu Ile Tyr Cys Phe
450 455 460
Leu Asn Gly Glu Val Arg Ala Val Leu Leu Lys Ser Leu Ala Thr Gln
465 470 475 480
Leu Ser Val Arg Gly His Pro Glu Trp Ala Pro Lys Arg Ala Ser Met
485 490 495

```
Tyr Ser Gly Ala Tyr Asn Thr Ala Pro Asp Thr Asp Ala Val Gln Pro
            500                 505                 510
Ala Gly Asp Pro Ser Ala Thr Gly Lys Arg Ile Ser Pro Pro Asn Lys
        515                 520                 525
Arg Leu Asn Gly Arg Lys Pro Ser Ser Ala Ser Ile Val Met Ile His
    530                 535                 540
Glu Pro Gln Gln Arg Gln Arg Leu Met Pro Arg Leu Gln Asn Lys Ala
545                 550                 555                 560
Arg Glu Lys Gly Lys Asp Arg Val Glu Lys Thr Asp Ala Glu Ala Glu
                565                 570                 575
Pro Asp Pro Thr Ile Ser His Ile His Ser Lys Glu Ala Gly Ser Ala
            580                 585                 590
Arg Ser Arg Thr Arg Gly Ser Lys Trp Ile Met Gly Ile Cys Phe Arg
        595                 600                 605
Gly Gln Lys Val Leu Arg Val Pro Ser Ala Ser Ser Val Pro Pro Glu
    610                 615                 620
Ser Val Val Phe Glu Leu Ser Glu Gln
625                 630


<210> 160
<211> 4385
<212> DNA
<213> Drosophila

<400> 160
gcttacaaaa ccaaaaacca gattttgtgc ttcaaatcac aaacataacg aaagaaatat 60
gaatggaaaa cagaggattt tatcatcact tgaaggctga caaaacattt ggcacaattt 120
acattgtccg gcgatttatt gccgtctaga gacaaaaaac actccggaaa attgtcactt 180
tgcttgcagt aagaaagtaa cttttaatta gaggttactg cagctatgca gctgctactg 240
cggttgctac taaaagcaaa agtgaagata tatcctttga atctcagcac aaattgcatt 300
ctccgtattt tctcatttgt ctaatggagt tatgtcacca agaaatgcgt gagtgttaca 360
ctcacaaaat atgttctatt cccaagaaat tcttctctgt aagtaagtgt aaaagtgtgt 420
acactaactc tgtttgtatt aaagacaaat gcataagaat gccttaaatt cgtctacgtg 480
taacatgtaa gccagaattt ttgttgcatg ttacacgttt tcccacaaat tgccaaaatc 540
atctgcttct gccggcccac tagtcgtatg cgtattattg attacggaac gaggccttgt 600
tttcttacca tttcggacaa atggtaatga catttcattg ttgattcgtt tgaaatcgta 660
cctttgacct gccctgaact tggacatcgt tctttgcatg tggcatgcaa tgccattgta 720
tttggggaat ttcgaagctt aaattctggt aaaattcaaa ccacaacact gaattcgttt 780
tttattcacc taacgcctga tgtttatttg caggtagcag gtagcacttt cgccacgcag 840
cctgcatatt tgagtgttgg aatttccgaa aagggaattt aaatgaaacg ccagaggagc 900
agcggggcgc actgagttgg cagccggaca ccgcggtcaa tcaaacggag tccttgcagg 960
acatcccacc accgactacc atctggcctt cattgacacc atggtcacga acatgtcgca 1020
gccgcattat tgcggcaccg gcatcgatga tttccacaca aagtgagtag aacgaggacg 1080
tggacttagg cctgggcaca tagcccctgg gtatgtctgc atattaatta gtttttttaag 1140
tgtaatttac ataatgattt gccagccggg gggcgcataa tgaaatcatg tgtgtaattg 1200
cctacggccg tggtggcaaa aaggctaacg aagtgtgcct aatgtggatg ggccccttttc 1260
ggattaaact gaacaaatac tcgcataaat ccataaagca gatgttaaag gtcggagatg 1320
aacattttga aaaatatatt gagttcaatc taaccaacca tgttgtttat tcggtttctt 1380
tctctacata agctctaatt actttaagaa tttaatcata cccaattttg tgctcaagtg 1440
aaactggttt taaatattta tcccaatttt ttttgtcttt ctccccaaca gctacaaata 1500
ctttcacggt tacttctcgc tgattgtctg tatcctggga accattgcga ataccctaaa 1560
tatcatagtg ctgacccgac gggagatgcg ctcccccaca aatgccatac tcacgggtct 1620
ggctgtcgcc gatctggctg tgatgctgga atacatacccc tatacggtgc acgactatat 1680
cctcagtgta aggctgccgc gagaggagca gctcagctac agctgggcgt gcttcatcaa 1740
gtttcactcg gtatttcccc aggtgctgca caccatctcc atttggctaa cggtgacgct 1800
ggcagtttgg cggtacatag cggtaagcta tccgcaaagg aatcgcatct ggtgtggaat 1860
gcgaactact ctgatcacca tagccacggc ctatgttgtc tgtgtcctgg tagtgtcacc 1920
ttggctgtac ctagtcacag ccattgccaa gttcctagag actttggatg ccaatggcaa 1980
gacgattgcc tcagtgccat tgagtcaata cattctggac tacaatcggc aggatgaggt 2040
```

gaccatgcag gtcatgtcga gtacaacgcc agatgtttcc tgggcgatac caagtgattc 2100
ggccaatgga actgcagtta gcttgctaag tctaaccaca gtgatacccc taaccacatt 2160
aagcactgga gtaaccacat cctcgtcgtt gggtgagcgc aatgtgactg tctataagct 2220
gtatcacagc gcactggcgc tgcgtgatcg gcagttcagg aatgcgacct tccttatata 2280
cagtgtcctg atcaagctga taccctgctt cgcactgacc attctgtctg tgcggctcat 2340
cggtgctctg ttggaggcca aaaggaggag gaagatcctg gcctgtcatg cagccaacga 2400
tatgcagcca attgtcaatg gaaaggtggt gattccgacg caacccaaga gctgtaaact 2460
gctggagaag gagaagcaga ccgatcgcac cacgaggatg cttctggcgg tactgctgct 2520
cttcctggtc accgagtttc cacagggcat tatgggtctg ctgaatgtgc tcctgggcga 2580
tgccttcttt ctgcaatgtt acctaaagct gagtaagtat tttaggatta agacacccac 2640
tcctccgctg tctgtaaccc gtaattgctt tattggctat gataattgac tgaatttatg 2700
gttggtgttt aggtgggccc agtttggcac attcatcaat catacgcacc gttggctgcc 2760
ataattgcac cacaaatttg aggcttcaat gaaggcctgg cattgacgtg ttcaataagt 2820
gtacattttg tttatctcgt acacccaaaa taaagtagtc ttacatttaa aacttcgcag 2880
atactattaa attgagagat ttttttcttt aatttgagta aaagtataaa attgaaaaaa 2940
aaatatttta aacatttaat atgtagtttt tagctagcat atgtttttttc aagtgtagga 3000
ggtataaaat ctaatcgccc ttgaacttta tgaaagccct gcccatagat ggcgtgttgg 3060
tggaacccga catttatggc atttcgattt tttcgtttca ggtgacctta tggacatctt 3120
ggcgcttatt aattcgagca tcaacttcat cctgtactgt tcgatgagcc gccagttccg 3180
gagcacgttc gcgctcctct tccgtccgcg ctggctggac aaatggctgc cgctgtcgca 3240
gcacgacggc gaagggaggg tgggcggaag tggcggcctg ggcggctacg gcggatatgg 3300
acggcagcgg ttgctgcaca cggatgccgt tagcaagagc atggccatcg atctcgggct 3360
gacgacccaa gtgacaaatg tgtagcagga gagcagcggc cgggcggcga tgtcagccgc 3420
agccggcgga gcagctgcat cggtggctct ggccctggca gccactgatg ttgatggatg 3480
tccgcctgcc accgatgctg ctgtttccac taacgacatc agcctggtcg agaagctaca 3540
tttgcagccc agtccaaggg ggactgcgat atccagtggc cagcatcgaa ggcggcgcag 3600
tgggagtggc accaagtgca tctggcccac gacggactgg ctgagaaagc tgcgtaacca 3660
gaaggctaga gaaacggagc aatcctccga acaggacata gagctgggca agagctccat 3720
caacaggcgc agcagtgttc ttctaatggt attactaagc agctccgatg aggttaaagc 3780
caaggcggtt ttggtcagtg aacagccgcc gagtccagcg gacgaggatg tggaagacgc 3840
cattgacgcc ctctggctgt gagacgaccc ctcacttcca ttttgtacgc cctgtaaatt 3900
atccaagtac ccactaatat ttaactaccg tgtacagacc aattattgag acaaattgtt 3960
atgcgctgat tccattattt tcccgccagt gttgtcttgg cgaatcgaga gccccttttcg 4020
ctgagacttt gctttgctgg aaaatgagtg gattttcctt gaaaggactc cctggaagca 4080
gcctgcaaca ttttgtgacc acatttgata agtgatatac ttgactaagt ttcccccctt 4140
aatagttgtg taagtgctaa gcgcgtctct agatataaac gaattagtca tagtctaact 4200
attaattaag atatattgat gaaataattc ggaagctaag agattaaatt tctttttata 4260
ttgaataaat gtttttttcga agtaaacttg attcaattca cttattgtaa gtgttattgc 4320
gacactttgt tttaattaaa cggaatttat attgaacttt gataaacgaa tgcttttgaa 4380
tgcat 4385

<210> 161
<211> 1389
<212> DNA
<213> Drosophila

<400> 161
atggtcacga acatgtcgca gccgcattat tgcggcaccg gcatcgatga tttccacaca 60
aactacaaat actttcacgg ttacttctcg ctgattgtct gtatcctggg aaccattgcg 120
aataccctaa atatcatagt gctgacccga cgggagatgc gctcccccac aaatgccata 180
ctcacgggtc tggctgtcgc cgatctggct gtgatgctgg aatacatacc ctatacggtg 240
cacgactata tcctcagtgt aaggctgccg cgagaggagc agctcagcta cagctgggcg 300
tgcttcatca agtttcactc ggtatttccc caggtgctgc acaccatctc catttggcta 360
acggtgacgc tggcagtttg gcggtacata gcggtaagct atccgcaaag gaatcgcatc 420
tggtgtggaa tgcgaactac tctgatcacc atagccacgg cctatgttgt ctgtgtcctg 480
gtagtgtcac cttggctgta cctagtcaca gccattgcca agttcctaga gactttggat 540
gccaatggca agacgattgc ctcagtgcca ttgagtcaat acattctgga ctacaatcgg 600
caggatgagg tgaccatgca ggtcatgtcg agtacaacgc cagatgtttc ctgggcgata 660
ccaataacca catcctcgtc gttgggtgag cgcaatgtga ctgtctataa gctgtatcac 720
agcgcactgg cgctgcgtga tcggcagttc aggaatgcga ccttccttat atacagtgtc 780

```
ctgatcaagc tgatacctg cttcgcactg accattctgt ctgtgcggct catcggtgct 840
ctgttggagg ccaaaaggag gaggaagatc ctggcctgtc atgcagccaa cgatatgcag 900
ccaattgtca atggaaaggt ggtgattccg acgcaaccca agagctgtaa actgctggag 960
aaggagaagc agaccgatcg caccacgagg atgcttctgg cggtactgct gctcttcctg 1020
gtcaccgagt ttccacaggg cattatgggt ctgctgaatg tgctcctggg cgatgccttc 1080
tttctgcaat gttacctaaa gctgagtgac cttatggaca tcttggcgct tattaattcg 1140
agcatcaact tcatcctgta ctgttcgatg agccgccagt tccggagcac gttcgcgctc 1200
ctcttccgtc cgcgctggct ggacaaatgg ctgccgctgt cgcagcacga cggcgaaggg 1260
agggtgggcg gaagtggcgg cctgggcggc tacggcggat atggacggca gcggttgctg 1320
cacacggatg ccgttagcaa gagcatggcc atcgatctcg ggctgacgac ccaagtgaca 1380
aatgtgtag                                                          1389
```

<210> 162
<211> 462
<212> PRT
<213> Drosophila

<400> 162

```
Met Val Thr Asn Met Ser Gln Pro His Tyr Cys Gly Thr Gly Ile Asp
1               5                   10                  15
Asp Phe His Thr Asn Tyr Lys Tyr Phe His Gly Tyr Phe Ser Leu Ile
            20                  25                  30
Val Cys Ile Leu Gly Thr Ile Ala Asn Thr Leu Asn Ile Ile Val Leu
        35                  40                  45
Thr Arg Arg Glu Met Arg Ser Pro Thr Asn Ala Ile Leu Thr Gly Leu
    50                  55                  60
Ala Val Ala Asp Leu Ala Val Met Leu Glu Tyr Ile Pro Tyr Thr Val
65                  70                  75                  80
His Asp Tyr Ile Leu Ser Val Arg Leu Pro Arg Glu Glu Gln Leu Ser
                85                  90                  95
Tyr Ser Trp Ala Cys Phe Ile Lys Phe His Ser Val Phe Pro Gln Val
            100                 105                 110
Leu His Thr Ile Ser Ile Trp Leu Thr Val Thr Leu Ala Val Trp Arg
        115                 120                 125
Tyr Ile Ala Val Ser Tyr Pro Gln Arg Asn Arg Ile Trp Cys Gly Met
    130                 135                 140
Arg Thr Thr Leu Ile Thr Ile Ala Thr Ala Tyr Val Val Cys Val Leu
145                 150                 155                 160
Val Val Ser Pro Trp Leu Tyr Leu Val Thr Ala Ile Ala Lys Phe Leu
                165                 170                 175
Glu Thr Leu Asp Ala Asn Gly Lys Thr Ile Ala Ser Val Pro Leu Ser
            180                 185                 190
Gln Tyr Ile Leu Asp Tyr Asn Arg Gln Asp Glu Val Thr Met Gln Val
        195                 200                 205
Met Ser Ser Thr Thr Pro Asp Val Ser Trp Ala Ile Pro Ile Thr Thr
    210                 215                 220
Ser Ser Ser Leu Gly Glu Arg Asn Val Thr Val Tyr Lys Leu Tyr His
225                 230                 235                 240
Ser Ala Leu Ala Leu Arg Asp Arg Gln Phe Arg Asn Ala Thr Phe Leu
                245                 250                 255
Ile Tyr Ser Val Leu Ile Lys Leu Ile Pro Cys Phe Ala Leu Thr Ile
            260                 265                 270
Leu Ser Val Arg Leu Ile Gly Ala Leu Leu Glu Ala Lys Arg Arg Arg
        275                 280                 285
Lys Ile Leu Ala Cys His Ala Ala Asn Asp Met Gln Pro Ile Val Asn
    290                 295                 300
Gly Lys Val Val Ile Pro Thr Gln Pro Lys Ser Cys Lys Leu Leu Glu
305                 310                 315                 320
Lys Glu Lys Gln Thr Asp Arg Thr Thr Arg Met Leu Leu Ala Val Leu
                325                 330                 335
```

```
Leu Leu Phe Leu Val Thr Glu Phe Pro Gln Gly Ile Met Gly Leu Leu
            340                 345                 350
Asn Val Leu Leu Gly Asp Ala Phe Phe Leu Gln Cys Tyr Leu Lys Leu
        355                 360                 365
Ser Asp Leu Met Asp Ile Leu Ala Leu Ile Asn Ser Ser Ile Asn Phe
    370                 375                 380
Ile Leu Tyr Cys Ser Met Ser Arg Gln Phe Arg Ser Thr Phe Ala Leu
385                 390                 395                 400
Leu Phe Arg Pro Arg Trp Leu Asp Lys Trp Leu Pro Leu Ser Gln His
                405                 410                 415
Asp Gly Glu Gly Arg Val Gly Gly Ser Gly Gly Leu Gly Gly Tyr Gly
            420                 425                 430
Gly Tyr Gly Arg Gln Arg Leu Leu His Thr Asp Ala Val Ser Lys Ser
        435                 440                 445
Met Ala Ile Asp Leu Gly Leu Thr Thr Gln Val Thr Asn Val
    450                 455                 460


<210> 163
<211> 4214
<212> DNA
<213> Drosophila

<400> 163
ataattctta tgaaaaatat aaaaatgggc tgattaaaat gattggtatt atggaaatta 60
ttattgaaaa aatatgaaat ttacatttta aaaagtgtta ttcgcgttat tactcagttt 120
gtaaatagcg ggcatcaatt gtgcatagtt tgaaggtaag ttgcccaata aatggcgcaa 180
agctcgtaac cgcgaattat atatgtatgt acatatacga taaaaataaa gtgtggcctt 240
tgcttggata aaaaaatcgt ccgagtcgcc ggatatccat attatcgctt tcagaaaatt 300
cgatatcatc gataggtgtt tccctttcaa aagtcttcgt ccgatttggt ccgtttgact 360
acttttatct ggtgcatacc cgatttgact gcttggacgg agtgggatgt gtgaatgaac 420
actaaaacgc ttgaaacttg agaaattagg gaagtagacg ggtagcgtca gtttgtttga 480
gaagattaac aagcccacaa actgcacaaa acttgcaagg ccattttatt atttgtctat 540
ccattttcga tcgatatggc aacaacattt aagccatgcc cactctaatg cgtcaataaa 600
tgcgactaat ccaggcgggc ccccaagtta acttggccct cctcatcgag tcaagcgctt 660
ccttcgaaga atgatttacc gttgatcata ccgggaagcc ttttttggtc ttcaggggag 720
gtagagacta catgtggtat taagatgaaa aaactaaatg cgtacatcaa atcaaaagta 780
ttcttattaa tccaagtgaa ttataatatt tattttggca ccttgtaaat gtgacaagtg 840
attagaaatg tggttcttct taaagatata aatgtacatc ataatatgct agagtactga 900
ggcactatta tttatatact ctcccgtgtg cgacagacta aatgattcat aatgaatatg 960
aataccagag ttcacagatg agcaataatc tcagttcctg ttaaagtgat gtcactgtcg 1020
gcttgtaggg agagtctgtg aaggaaaata tcaatgttaa aagaagataa attaaagcaa 1080
taagatgaac accaacctga agattttaga tcgcctttct catcgaactt gtatacgcga 1140
gcattaataa tactcagatc tgtctgactt gattccctgt aaaaaaaaaa aatgggcaat 1200
taaaatatta tagcttctat tcgaatattt caacttacac attaatggcc caccaggatt 1260
tgacgtgcaa attcttgcca aggagcagaa taaaaccagc tatggccagg cccagaatgc 1320
aaattccact ccagaagtca atgatgaaga aatgctctct cactctggcc attatataca 1380
cgactatttc tcggatccag accgcaaaca cgcccactat gtaagcgcta aaggcataga 1440
aactgcagga atataaaaga taaataggtc tataataaaa acagctttta ggaaatctta 1500
ctgcgtcttg acgattttca gttgagtgcc caatatgtac tgagtatctg gtggcagttc 1560
gtagatggaa aagtacgcct tgagacagca tacaatgctg aatatcagga ggagtgcaat 1620
gggtccgtag tagtagatca aaataccccca ggtgcgaact gaaaatgcat agataattcg 1680
aaattagact ttgttgtact ttcttattca acctaccatc gaaccagcag taatcttctc 1740
ccatgcctgg cttgaagtgc ttgggcagct tggagtcctg ggctaaggag accaaaaacc 1800
gaagaccaat caccgccaga actgccagag caagggacaa tcgacggacg gagctcctga 1860
cagccttttg tttaaatttg agcagaaagt caagactgca aatggccaga aagacgaagg 1920
agagcatgat gcagaagtag gccagaaatc cttttgggat gcatagaaca aaataagtta 1980
tccagttgaa gtagacaaaa ttactcaacc tatgtttctg caagccacat gcgagagctt 2040
catgggattc ttcaaggcca agtagaccaa cagggagtaa ccaacgatca tggacagcaa 2100
gtagtagatg attaactggc catagtggct tttacgggca tctctcacag atccaagtag 2160
```

ggagagtatg aaaatgttta taatgatggc tataatactg cctagaaggg aacaagagtt 2220
agaatctctt cgaaagtatt caaacaacac gactcacaaa tggcgtaaat ccaaacccgg 2280
taccccgttt gaaagcgttc gcagttgagc ggagtgagct cccactgctc cggattgtgc 2340
tccaacgggc taaagcaata ctcatccgtt gaccacagac gattgtcccg ccgcagagtt 2400
ccattctcga agaggtccca ttgccagaag ttatcgtgct ttttgtccac gaatttgttc 2460
cggcaaccca gctctgtgcg caccacaaaa cgatcgcgta tgcgtacttg atccacagtc 2520
ctgttagcgt aggtgagctc cacgtgagtg gactcgcgga cttggtgtgg cattgtgcag 2580
ttccagtgct tggcatcaaa gacgaggttc tccggacaac agaaactgat gcagggcttc 2640
agcagacaga cgcagccgcg cagatgtttc ttggcccggt actccactcc atcgatcacc 2700
ttgaaggagt actccgccat cagatgcggt gggaccacta ctccagcgta tgagtaggag 2760
ccatccttca tccgcagtcc atccgtaatg ttgaccgtgt gggcgtaggc acagggatgg 2820
tttatactgg cgatctccac tgactttgcg cccacggaga gccaaataat taaaagaatt 2880
atcaatggcg agaccatggc taactgatcc acgcttgttg gatttttggt cgattcgttc 2940
gtcggcctga cgtcacacgc tgttcagctc gataattaag taaacaaata tgagaagcga 3000
tcgccgttcg ctttgagcaa tgcacgcaca ctcgccgagt cagaaatctg caccaacaac 3060
aatggaaacc caaaattgac aacgaatgtg acgcgggtct aaaaacggct gcgattgcga 3120
caacgacacc gacaacacca aaaagaacaa ccataaacat aaaagtttgg tggagagtct 3180
ctcctagcta gtcgtcttgc ctctctctaa acgacaaacc ggcaaccgca acggagctcc 3240
atgctggcca agaataaaaa acagactaac tgattagccg ttcccaaaac cagacgcgac 3300
cacaagcagc tggcaaaaag agcgattggg aagactcaca acccgtcccg cccctcctcc 3360
tagaaaccgg cgtggaatct caaccagctc ggaacttgca catgcatacc ttaagttgta 3420
agcatataat ccattcgagc catttccaga cttatcttga gatagatata aaagttaaga 3480
atattagatt agattatttc catccatttt ctataattta ttgcctttgc gaccgtaaag 3540
aaaatccaat ttttaaccga caatccaaat ggcccaaacg aatttaccac agctttagtt 3600
tttaaatatt taggtgcatt tccaggctat acaaagtttg tctacatgta ggtgtgtgac 3660
ctggcttgat atcgtttacg gtcttcaaga atattcgaaa agattcgttg atttttgaca 3720
aacagatccg ctgaggaaaa aacgacgttc attttatctt tggctttatt aagcttgaaa 3780
tacaattata attcttgtat tgtgttttac attagttcag agataatgca gatatggaga 3840
atggtacatt tatgcaaaga gggtatacaa tcaacaccaa tctcgttttt aacaaggcac 3900
atctgtatgt tacgtatttt tgctggatgt ttaccttgga cttaggctaa tggagaatta 3960
caaatttgat ctctcagctg ggttcggaga accttacaat tttgcttgca taaacagtag 4020
atatggctct gcataaattt cttatacaac taatcaggtt gttcgtggtc acatggcatt 4080
aaaaaatgaa agagaattgg gcggggtggg gtgggtaagg gttttcgtgg tagttctggg 4140
ggatagggac tttctcttat agtagtgtct ttcataaagt tctatccgta tgtggatgct 4200
acttcgcgaa ttta 4214

<210> 164
<211> 1859
<212> DNA
<213> Drosophila

<400> 164
tcgtttagag agaggcaaga cgactagcta ggagagactc tccaccaaac ttttatgttt 60
atggttgttc tttttggtgt tgtcggtgtc gttgtcgcaa tcgcagccgt ttttagaccc 120
gcgtcacatt cgttgtcaat tttgggtttc cattgttgtt ggtgcagatt tctgactcgg 180
cgagtgtgcg tgcattgctc aaagcgaacg gcgatcgctt ctcatatttg tttacttaat 240
tatcgagctg aacagcgtgt gacgtcaggc cgacgaacga atcgaccaaa aatccaacaa 300
gcgtggatca gttagccatg gtctcgccat tgataattct tttaattatt tggctctccg 360
tgggcgcaaa gtcagtggag atcgccagta taaaccatcc ctgtgcctac gcccacacgg 420
tcaacattac ggatggactg cggatgaagg atggctccta ctcatacgct ggagtagtgg 480
tcccaccgca tctgatggcg gagtactcct tcaaggtgat cgatggagtg gagtaccggg 540
ccaagaaaca tctgcgcggc tgcgtctgtc tgctgaagcc ctgcatcagt ttctgttgtc 600
cggagaacct cgtctttgat gccaagcact ggaactgcac aatgccacac caagtccgcg 660
agtccactca cgtggagctc acctacgcta acaggactgt ggatcaagta cgcatacgcg 720
atcgttttgt ggtgcgcaca gagctgggtt gccggaacaa attcgtggac aaaaagcacg 780
ataacttctg gcaatgggac ctcttcgaga atggaactct gcggcgggac aatcgtctgt 840
ggtcaacgga tgagtattgc tttagcccgt tggagcacaa tccggagcag tgggagctca 900
ctccgctcaa ctgcgaacgc tttcaaacgg ggtaccgggt ttggatttac gccatttgca 960
gtattatagc catcattata aacattttca tactctccct acttggatct gtgagagatg 1020
cccgtaaaag ccactatggc cagttaatca tctactactt gctgtccatg atcgttggtt 1080

actccctgtt ggtctacttg gccttgaaga atcccatgaa gctctcgcat gtggcttgca 1140
gaaacatagg atttctggcc tacttctgca tcatgctctc cttcgtcttt ctggccattt 1200
gcagtcttga ctttctgctc aaatttaaac aaaaggctgt caggagctcc gtccgtcgat 1260
tgtcccttgc tctggcagtt ctggcggtga ttggtcttcg gtttttggtc tccttagccc 1320
aggactccaa gctgcccaag cacttcaagc caggcatggg agaagattac tgctggttcg 1380
atgttcgcac ctggggtatt ttgatctact actacggacc cattgcactc ctcctgatat 1440
tcagcattgt atgctgtctc aaggcgtact tttccatcta cgaactgcca ccagatactc 1500
agtacatatt gggcactcaa ctgaaaatcg tcaagacgca tttctatgcc tttagcgctt 1560
acatagtggg cgtgtttgcg gtctggatcc gagaaatagt cgtgtatata atggccagag 1620
tgagagagca tttcttcatc attgacttct ggagtggaat ttgcattctg ggcctggcca 1680
tagctggttt tattctgctc cttggcaaga atttgcacgt caaatcctgg tgggccatta 1740
atgtggaatc aagtcagaca gatctgagta ttattaatgc tcgcgtatac aagttcgatg 1800
agaaaggcga tctaaaatct tcagactctc cctacaagcc gacagtgaca tcactttaa 1859

<210> 165
<211> 513
<212> PRT
<213> Drosophila

<400> 165
Met Val Ser Pro Leu Ile Ile Leu Leu Ile Ile Trp Leu Ser Val Gly
1 5 10 15
Ala Lys Ser Val Glu Ile Ala Ser Ile Asn His Pro Cys Ala Tyr Ala
20 25 30
His Thr Val Asn Ile Thr Asp Gly Leu Arg Met Lys Asp Gly Ser Tyr
35 40 45
Ser Tyr Ala Gly Val Val Val Pro Pro His Leu Met Ala Glu Tyr Ser
50 55 60
Phe Lys Val Ile Asp Gly Val Glu Tyr Arg Ala Lys Lys His Leu Arg
65 70 75 80
Gly Cys Val Cys Leu Leu Lys Pro Cys Ile Ser Phe Cys Cys Pro Glu
85 90 95
Asn Leu Val Phe Asp Ala Lys His Trp Asn Cys Thr Met Pro His Gln
100 105 110
Val Arg Glu Ser Thr His Val Glu Leu Thr Tyr Ala Asn Arg Thr Val
115 120 125
Asp Gln Val Arg Ile Arg Asp Arg Phe Val Val Arg Thr Glu Leu Gly
130 135 140
Cys Arg Asn Lys Phe Val Asp Lys Lys His Asp Asn Phe Trp Gln Trp
145 150 155 160
Asp Leu Phe Glu Asn Gly Thr Leu Arg Arg Asp Asn Arg Leu Trp Ser
165 170 175
Thr Asp Glu Tyr Cys Phe Ser Pro Leu Glu His Asn Pro Glu Gln Trp
180 185 190
Glu Leu Thr Pro Leu Asn Cys Glu Arg Phe Gln Thr Gly Tyr Arg Val
195 200 205
Trp Ile Tyr Ala Ile Cys Ser Ile Ile Ala Ile Ile Ile Asn Ile Phe
210 215 220
Ile Leu Ser Leu Leu Gly Ser Val Arg Asp Ala Arg Lys Ser His Tyr
225 230 235 240
Gly Gln Leu Ile Ile Tyr Tyr Leu Leu Ser Met Ile Val Gly Tyr Ser
245 250 255
Leu Leu Val Tyr Leu Ala Leu Lys Asn Pro Met Lys Leu Ser His Val
260 265 270
Ala Cys Arg Asn Ile Gly Phe Leu Ala Tyr Phe Cys Ile Met Leu Ser
275 280 285
Phe Val Phe Leu Ala Ile Cys Ser Leu Asp Phe Leu Leu Lys Phe Lys
290 295 300
Gln Lys Ala Val Arg Ser Ser Val Arg Arg Leu Ser Leu Ala Leu Ala
305 310 315 320

```
Val Leu Ala Val Ile Gly Leu Arg Phe Leu Val Ser Leu Ala Gln Asp
                325                 330                 335
Ser Lys Leu Pro Lys His Phe Lys Pro Gly Met Gly Glu Asp Tyr Cys
            340                 345                 350
Trp Phe Asp Val Arg Thr Trp Gly Ile Leu Ile Tyr Tyr Tyr Gly Pro
        355                 360                 365
Ile Ala Leu Leu Leu Ile Phe Ser Ile Val Cys Cys Leu Lys Ala Tyr
    370                 375                 380
Phe Ser Ile Tyr Glu Leu Pro Pro Asp Thr Gln Tyr Ile Leu Gly Thr
385                 390                 395                 400
Gln Leu Lys Ile Val Lys Thr His Phe Tyr Ala Phe Ser Ala Tyr Ile
                405                 410                 415
Val Gly Val Phe Ala Val Trp Ile Arg Glu Ile Val Val Tyr Ile Met
            420                 425                 430
Ala Arg Val Arg Glu His Phe Phe Ile Ile Asp Phe Trp Ser Gly Ile
        435                 440                 445
Cys Ile Leu Gly Leu Ala Ile Ala Gly Phe Ile Leu Leu Leu Gly Lys
    450                 455                 460
Asn Leu His Val Lys Ser Trp Trp Ala Ile Asn Val Glu Ser Ser Gln
465                 470                 475                 480
Thr Asp Leu Ser Ile Ile Asn Ala Arg Val Tyr Lys Phe Asp Glu Lys
                485                 490                 495
Gly Asp Leu Lys Ser Ser Asp Ser Pro Tyr Lys Pro Thr Val Thr Ser
            500                 505                 510
Leu
```

```
<210> 166
<211> 5208
<212> DNA
<213> Drosophila

<400> 166
cgagccattt ccagacttat cttgagatag atataaaagt taagaatatt agattagatt 60
atttccatcc atttctata atttattgcc tttgcgaccg taaagaaaat ccaattttta 120
accgacaatc caaatggccc aaacgaattt accacagctt tagtttttaa atatttaggt 180
gcatttccag gctatacaaa gtttgtctac atgtaggtgt gtgacctggc ttgatatcgt 240
ttacggtctt caagaatatt cgaaaagatt cgttgatttt tgacaaacag atccgctgag 300
gaaaaaacga cgttcatttt atctttggct ttattaagct tgaaatacaa ttataattct 360
tgtattgtgt tttacattag ttcagagata atgcagatat ggagaatggt acatttatgc 420
aaagagggta tacaatcaac accaatctcg tttttaacaa ggcacatctg tatgttacgt 480
atttttgctg gatgtttacc ttggacttag gctaatggag aattacaaat ttgatctctc 540
agctgggttc ggagaacctt acaattttgc ttgcataaac agtagatatg gctctgcata 600
aatttcttat acaactaatc aggttgttcg tggtcacatg gcattaaaaa atgaaagaga 660
attgggcggg gtggggtggg taagggtttt cgtggtagtt ctgggggata gggactttct 720
cttatagtag tgtctttcat aaagttctat ccgtatgtgg atgctacttc gcgaatttat 780
agatatcgac tttaaatgtt gattacagaa taatctcatg ataggaatat gtagtcctaa 840
aatatgaata agtgatcttt atctacagaa aggaacgcag tgttcgcaat aatgcgagga 900
tccgatcgaa atgaatataa atgcgattag tccttcctga tctgcacgca tcccgaggta 960
ttcaatgctc tgcctatggc ggcaactgcg agtgctgtga tcagcggaag atggtggttt 1020
tctgaggatc cacgctcgag ctgatcacta gcggcttctc caccgatggc ttctcgtgca 1080
gcgacagatt ggccacgctg ctcgacgtgg tcttggtgga gtactggctc tgacgttggt 1140
tgctgccatc gcgaacactc gagcatctgt gtttgtgtga agtggtggag tggttatgac 1200
gtccattgcg tgggcagcgt ggaatagaca atggcggtaa atctgattag aattggttcg 1260
tttgagtgct tggattttaa gttagggaaa agcagacaca ttaagcgatc ctacaggata 1320
taaacaggtt accttcacca agacacaaaa gtggatggtc attatgatta gttttcgagg 1380
ggttagaggt tatcagggtt atctaatggt ttcttccaac tggggttaac tacaatgtca 1440
actacttctc catttccaga ggattacctg tgaccatttt gattagatgg ctatcttcgg 1500
ggtcggactc gggtcggtgg cattgctacc ggccttgtca acagccttaa ggtgcagctt 1560
```

tcttggcagc acattcgcgt cagaagtgat ggaagtcgag gtagtggtcg agctcaggcc 1620
attcgagcgc cttttctcca ggagtcttca atgaatagcg atcgatgatg atcaaattga 1680
tgttgagaca catgcacata tggacacaca tggtggatgc aaacataagt ttgtgtttag 1740
atcagggcaa ttaggaatgg tacttctgtt tagtagctac aaagacgcga ccggactgag 1800
tcaccgagct aggcggctgg agattcagtg ctgtgttggc tgtcagtgtt cggtttagag 1860
ttagtctgca ggcagtgtct gttggtgttg agcttgcttt taagcaatat cgaaagcagt 1920
ctgacaacga aaataagtac aacatataag gtttaaaata atcacagatt gcgcgctgtc 1980
agaactcgaa gtgcatatgt attcccataa tcttcaatgg gtcttccatc ttagtgcaag 2040
agatttaagg aacgtggaat tagacaggaa tcatataacc cgaaattttc aatattcttg 2100
atttaagcgc gggttagcaa aggccacaac aaacgagcaa ccgagctaaa ggagaccttc 2160
ttactgaaat cgacgaaatc gatttagact tcgtgtgaat gaataccact ggtggaccca 2220
ttggcaaatg cgttgcggat caccaccagc ttgcggatct ggttggaggc gtccataaac 2280
cgctgcttca tcgagaacga agtcggtccc atcgttacgc tctcagtctg ccgcagactg 2340
cgcacacgta tcgttctgag aatagtcggg gttacagata tgatggtgga cagagtcaga 2400
gacacgacac acaagaggga gtcacgagtc gtgtgaatgg acaaacgccc agtgggcggg 2460
acaggacata acgtggagtt aagaatagag caacacttac ctgtttgtga tcaagtgctt 2520
gaccttcttc ttcattacaa acaacatgaa aatgagaaag ccctgcatcg cattggccag 2580
atccgagata taaaagagct tagaccatcc cttgtcgctg cccacgaagt acgagataag 2640
ctccgtgagc caggtgatgc ccattatgag aaagaggcgt aagaacaagc cgaatctata 2700
gttggaccaa gcacgataaa aacgtcttta ggggaaggtg gataagaaat actgaaagga 2760
tgacactcac ttgtccttct cggttcgtag gttcttggtg ctgttctcac tcgcaattat 2820
ccgggccatc tcgcgctgca cgccgtgaat cttaatcgcc gtcattatga acatgatggt 2880
gttggcaaca acgatcgcca gaatgggacc gtagaaatag atcatggctg cccagttgct 2940
catatccagc cagcagtata caccatcacc aatacctggc tttaggttcg cgggcaaatt 3000
ggtaagttgc tgggcgcaat aggtgaaggc taggaaaact agggcaatac cccaagagta 3060
cagcgaataa aacaggaagc gtttcttctc ctggaatcgg ttaattcctc tcgtccctcg 3120
gaaattgtgc cacaagtcaa agctaatgac acttagccac atgtaggcac ccataaaaaa 3180
gaaatatgct gtgtatccaa acactttgca tgtaacacca gtggcatcaa cttgatagag 3240
ttgcacgtag cacagcgagg agtaacccac ggagagccct attaaatagc agaccaagct 3300
ctttccgtgc tggttgcgga gctcgggaat caataggtaa acggctatgg taagcatcat 3360
gaagggaatg gagaacatca tggctagtgc aaggaatttc ggtcggatta ataaaatgct 3420
gtttacaaga agtaaaggtg atcgtttcac ttaccatacg aattgataat ctttaccgtg 3480
ctatgatcgt cctgcatgtc acaattggcc ggatgtatgg tgtagaacat gtccgattcg 3540
ttgacgtatg tcggaaccat gcagaactcg ttcttcctga ttagcagttt atcgtccaca 3600
cggagcatgg atccgttctc aaagagtaga taatcatccc aattgttggt ctctggctgc 3660
agagaaaaca tcttggggca gggtctgaag ctttggatag caaactgcgt aaacatgttg 3720
accaatatcg tagactcgtt gcggaaggtt atattgagca taggcggatc tggccacgtt 3780
cggttgtccg agtggtctat tatgcatccg tcaacttcag agttaaagac ctgccgccac 3840
gggcagcaga tattcaggca ggacttgaac ttgcagacgc atcccctaac gtgttcccga 3900
acctctatgc gctccgttag cgagctgtgt atgtagtcgt acgtcccgac caattctgcg 3960
ggaactattg tgccgtagta atcatagctt ccgttgggaa agagcctatg gcccgtcaga 4020
ttcactgtat cgtagaaatt acagggcaca tttggatcgt ctatggtgtg gtagtggtct 4080
ctctcggcca taacgtaagt gccgaaggga atcccggaa tcaaccggaa cacgaccaag 4140
cacaggacgc ccagcaatgt aactccgcag tatagggatc ctccaggctg atgaattttc 4200
ttcggcataa tgaaggttct gaaattcgta aatctttatt aaatattatt acaaatattt 4260
gttattattc tttaacgttt attcttttaa cgtttaaata ttcagtggaa tatagttatt 4320
acagttagtc atcagtcgca acacgcagca attaagcctt agcaagtgaa gtcactaatt 4380
aagacctatt tgccgattag ttagtcgttt aatgtttatc taagtgagtt cactaatttg 4440
ggcaatagag caacatctag ggacccatta gcggggacga gtggtgaatg atctatgcgg 4500
acaacaggtt gtcgaaaaga attaaggagg cgaaacagga tagcgtcaca agcaattagc 4560
agccaacgtg tcccatgggg aggaaccagg ggcggactac ccacgacctg agcgccattc 4620
tccgcgggac atcaatcata aggtcaacca cgcagggagt gcagactgag caaacattta 4680
tggattcgtg ttaaattgat agtaactgcc attgtctgtc aatagaggtg acagtctccc 4740
cgaactccac tccaaaccaa tcccgatctg cccacttttc aagcgctgct aacgagcccg 4800
tccataaatt cattaccagt gcgaaaagtt ataccgctcg ctatcttatc ttgtatcatt 4860
agcctttggg gaccaactaa ttcattgatc cgctgattgt ttgcctttgg gggagaagca 4920
gaagggtaac ctgcggcatt gtcaggagtg tgggccacta aattagtccg aatttatggg 4980
atgcgaacgg gtggagaagg acaaacacat taaactccgc acagttggct aatggggaaa 5040
atgaatactg gcagtgggtg gggaagggag agccagagca cttatcttgc ggtgggaaat 5100
tagcataatt aatcgctgag cacatgcact tgggaaattt cagcggacca aattcaatcg 5160

tatatatctg cgcactttaa atgtgcgaat gagtgctgcg gatgtttg 5208

<210> 167
<211> 1914
<212> DNA
<213> Drosophila

<400> 167
atgccgaaga aaattcatca gcctggagga tccctatact gcggagttac attgctgggc 60
gtcctgtgct tggtcgtgtt ccggttgatt ccggggattc ccttcggcac ttacgttatg 120
gccgagagag accactacca caccatagac gatccaaatg tgccctgtaa tttctacgat 180
acagtgaatc tgacgggcca taggctcttt cccaacggaa gctatgatta ctacggcaca 240
atagttcccg cagaattggt cgggacgtac gactacatac acagctcgct aacggagcgc 300
atagaggttc gggaacacgt taggggatgc gtctgcaagt tcaagtcctg cctgaatatc 360
tgctgcccgt ggcggcaggt ctttaactct gaagttgacg gatgcataat agaccactcg 420
gacaaccgaa cgtggccaga tccgcctatg ctcaatataa ccttccgcaa cgagtctacg 480
atattggtca acatgtttac gcagtttgct atccaaagct tcagaccctg ccccaagatg 540
ttttctctgc agccagagac caacaattgg gatgattatc tactctttga gaacggatcc 600
atgctccgtg tggacgataa actgctaatc aggaagaacg agttctgcat ggttccgaca 660
tacgtcaacg aatcggacat gttctacacc atacatccgg ccaattgtga catgcaggac 720
gatcatagca cggtaaagat tatcaattcg tatgccgttt acctattgat tcccgagctc 780
cgcaaccagc acggaaagag cttggtctgc tatttaatag ggctctccgt gggttactcc 840
tcgctgtgct acgtgcaact ctatcaagtt gatgccactg gtgttacatg caaagtgttt 900
ggatacacag catatttctt ttttatgggt gcctacatgt ggctaagtgt cattagcttt 960
gacttgtggc acaatttccg agggacgaga ggaattaacc gattccagga gaagaaacgc 1020
ttcctgtttt attcgctgta ctcttggggt attgccctag ttttcctagc cttcacctat 1080
tgcgcccagc aacttaccaa tttgcccgcg aacctaaagc caggtattgg tgatggtgta 1140
tactgctggc tggatatgag caactgggca gccatgatct atttctacgg tcccattctg 1200
gcgatcgttg ttgccaacac catcatgttc ataatgacgg cgattaagat tcacggcgtg 1260
cagcgcgaga tggcccggat aattgcgagt gagaacagca ccaagaacct acgaaccgag 1320
aaggacaaat tcggcttgtt cttacgcctc tttctcataa tgggcatcac ctggctcacg 1380
gagcttatct cgtacttcgt gggcagcgac aagggatggt ctaagctctt ttatatctcg 1440
gatctggcca atgcgatgca gggctttctc attttcatgt tgtttgtaat gaagaagaag 1500
gtcaagcact tgatcacaaa cagaacgata cgtgtgcgca gtctgcggca gactgagagc 1560
gtaacgatgg gaccgacttc gttctcgatg aagcagcggt ttatggacgc ctccaaccag 1620
atccgcaagc tggtggtgat ccgcaacgca tttgccaatg ggtccaccag tgatttaccg 1680
ccattgtcta ttccacgctg cccacgcaat ggacgtcata accactccac cacttcacac 1740
aaacacagat gctcgagtgt tcgcgatggc agcaaccaac gtcagagcca gtactccacc 1800
aagaccacgt cgagcagcgt ggccaatctg tcgctgcacg agaagccatc ggtggagaag 1860
ccgctagtga tcagctcgag cgtggatcct cagaaaacca ccatcttccg ctga 1914

<210> 168
<211> 637
<212> PRT
<213> Drosophila

<400> 168
Met Pro Lys Lys Ile His Gln Pro Gly Gly Ser Leu Tyr Cys Gly Val
1 5 10 15
Thr Leu Leu Gly Val Leu Cys Leu Val Val Phe Arg Leu Ile Pro Gly
20 25 30
Ile Pro Phe Gly Thr Tyr Val Met Ala Glu Arg Asp His Tyr His Thr
35 40 45
Ile Asp Asp Pro Asn Val Pro Cys Asn Phe Tyr Asp Thr Val Asn Leu
50 55 60
Thr Gly His Arg Leu Phe Pro Asn Gly Ser Tyr Asp Tyr Tyr Gly Thr
65 70 75 80
Ile Val Pro Ala Glu Leu Val Gly Thr Tyr Asp Tyr Ile His Ser Ser

```
                85                  90                  95
Leu Thr Glu Arg Ile Glu Val Arg Glu His Val Arg Gly Cys Val Cys
            100                 105                 110
Lys Phe Lys Ser Cys Leu Asn Ile Cys Cys Pro Trp Arg Gln Val Phe
        115                 120                 125
Asn Ser Glu Val Asp Gly Cys Ile Ile Asp His Ser Asp Asn Arg Thr
    130                 135                 140
Trp Pro Asp Pro Pro Met Leu Asn Ile Thr Phe Arg Asn Glu Ser Thr
145                 150                 155                 160
Ile Leu Val Asn Met Phe Thr Gln Phe Ala Ile Gln Ser Phe Arg Pro
                165                 170                 175
Cys Pro Lys Met Phe Ser Leu Gln Pro Glu Thr Asn Asn Trp Asp Asp
            180                 185                 190
Tyr Leu Leu Phe Glu Asn Gly Ser Met Leu Arg Val Asp Asp Lys Leu
        195                 200                 205
Leu Ile Arg Lys Asn Glu Phe Cys Met Val Pro Thr Tyr Val Asn Glu
    210                 215                 220
Ser Asp Met Phe Tyr Thr Ile His Pro Ala Asn Cys Asp Met Gln Asp
225                 230                 235                 240
Asp His Ser Thr Val Lys Ile Ile Asn Ser Tyr Ala Val Tyr Leu Leu
                245                 250                 255
Ile Pro Glu Leu Arg Asn Gln His Gly Lys Ser Leu Val Cys Tyr Leu
            260                 265                 270
Ile Gly Leu Ser Val Gly Tyr Ser Ser Leu Cys Tyr Val Gln Leu Tyr
        275                 280                 285
Gln Val Asp Ala Thr Gly Val Thr Cys Lys Val Phe Gly Tyr Thr Ala
    290                 295                 300
Tyr Phe Phe Phe Met Gly Ala Tyr Met Trp Leu Ser Val Ile Ser Phe
305                 310                 315                 320
Asp Leu Trp His Asn Phe Arg Gly Thr Arg Gly Ile Asn Arg Phe Gln
                325                 330                 335
Glu Lys Lys Arg Phe Leu Phe Tyr Ser Leu Tyr Ser Trp Gly Ile Ala
            340                 345                 350
Leu Val Phe Leu Ala Phe Thr Tyr Cys Ala Gln Gln Leu Thr Asn Leu
        355                 360                 365
Pro Ala Asn Leu Lys Pro Gly Ile Gly Asp Gly Val Tyr Cys Trp Leu
    370                 375                 380
Asp Met Ser Asn Trp Ala Ala Met Ile Tyr Phe Tyr Gly Pro Ile Leu
385                 390                 395                 400
Ala Ile Val Val Ala Asn Thr Ile Met Phe Ile Met Thr Ala Ile Lys
                405                 410                 415
Ile His Gly Val Gln Arg Glu Met Ala Arg Ile Ile Ala Ser Glu Asn
            420                 425                 430
Ser Thr Lys Asn Leu Arg Thr Glu Lys Asp Lys Phe Gly Leu Phe Leu
        435                 440                 445
Arg Leu Phe Leu Ile Met Gly Ile Thr Trp Leu Thr Glu Leu Ile Ser
    450                 455                 460
Tyr Phe Val Gly Ser Asp Lys Gly Trp Ser Lys Leu Phe Tyr Ile Ser
465                 470                 475                 480
Asp Leu Ala Asn Ala Met Gln Gly Phe Leu Ile Phe Met Leu Phe Val
                485                 490                 495
Met Lys Lys Lys Val Lys His Leu Ile Thr Asn Arg Thr Ile Arg Val
            500                 505                 510
Arg Ser Leu Arg Gln Thr Glu Ser Val Thr Met Gly Pro Thr Ser Phe
        515                 520                 525
Ser Met Lys Gln Arg Phe Met Asp Ala Ser Asn Gln Ile Arg Lys Leu
    530                 535                 540
Val Val Ile Arg Asn Ala Phe Ala Asn Gly Ser Thr Ser Asp Leu Pro
545                 550                 555                 560
Pro Leu Ser Ile Pro Arg Cys Pro Arg Asn Gly Arg His Asn His Ser
```

```
                565                 570                 575
Thr Thr Ser His Lys His Arg Cys Ser Ser Val Arg Asp Gly Ser Asn
            580                 585                 590
Gln Arg Gln Ser Gln Tyr Ser Thr Lys Thr Thr Ser Ser Ser Val Ala
        595                 600                 605
Asn Leu Ser Leu His Glu Lys Pro Ser Val Glu Lys Pro Leu Val Ile
    610                 615                 620
Ser Ser Ser Val Asp Pro Gln Lys Thr Thr Ile Phe Arg
625                 630                 635


<210> 169
<211> 3957
<212> DNA
<213> Drosophila

<400> 169
ccagacatgc ttcctcgatg ggcttaacg tcagcggagc agcgatgtca gtggcagcca 60
tcatctgaac cgaatgacag tggaggtcgc tagaaaaggg gttccatttc cactagttta 120
tactaccgaa ctgatactta aatcaaggca gttaaaaatt taatatgaaa ccggaaatgg 180
cttcaacatc tcgttaacgc tctgcccaat tctacagtca tccgatttta atgaagcaaa 240
atttactgcc atgccatgca agagatatga atcaccgcat gggaatagcc cacaaaggac 300
ccttacccca tgcactctat catgaacacc gcacacactt ttaatctgcg actaattgct 360
gctccgtatt gaattgattt cggttgcccc aaagtgccac attccaaatc tccgttgatt 420
ggggaaatcc ccgaaaaaca ttcgaaccat tcgtccctgg cgagcacaga aacagcgtct 480
ctgattctgg caggaattgg agggcatagg gcggaccttg tggcggtttg tcgagtgcac 540
aaagtgcaag gttgcgtggc cattcatttc gcagccgatg ggtggaagtg ggcagcggtg 600
ggtggcgccc ggaaggaaga gcatcagtaa tctctaagtg cgccttgttc gcattcgtac 660
tcgtgccagc attgcataag tgtcagtgtc aacggggcgg ggatgtggtt gtgcttggtg 720
gttatgtggg tgtttgtgtg gtagtagtgt ccttgttggc agacaaaaag actgatgatg 780
atggcatttt gttactttgc taccttcctt tgtcgcccgt cgcacggcag ctcctttagt 840
atgaaatgag gcgaccaggc aagcgcctcg tcatcgaaga tcacccaaac agagtacgga 900
ccaggacacg gcaactcgca caggcacagg aacaggtgca ggcacaacaa ctggcaggag 960
catcggtgtc ggcaagggat actggcgcgg gatcgaaagc atgaacaggg acaatctgca 1020
gcagtggtgg gagaactcgt atcggcgcca acatccagag cccaccgatg acctcggcct 1080
ggactccgcg gagctgcatc tcgccctcca ggagcccaac cagctgcccg cggactacga 1140
ctatggcaac tttagtctag gcaatccgta tgacgtggac tcagagcact ccatttcccc 1200
gctgacgctg ctcctgctcg ctgttagcta tggcctggtg gttttcggcg gcgtcgtggg 1260
caactccacg ctcgttctga ccctctgctc cgcctcttcg gtgcgtttgc gtaacccgct 1320
gctgctggcc gtctgcattg ccgatctcct ggtcaccggc atctccgctc cggtcacgct 1380
cctcaatctc gcgatgaacc ggaggacgcg atcgctgccc cttgtgctgt gcaaggtcat 1440
acactacgta caggtgagtg ggtcaatcaa acaagcttaa gttagccaag taactaagtt 1500
aaccccacgc tattcaggtc atgcccgttt cggccagcac tatttcgttt ttcatgctct 1560
cccttgaccg ctacgccaca gtaaagcatc cgcgattggc gcagcttcgc cagcggcgct 1620
acttgcacgt ctccctggcg ctgctatcct ggctcgcctc tgccgccatc agcactccct 1680
ttctgtttgc ctacaaaatc atcgccaagt cgatggtggt caagggcggt ggggcggcga 1740
acaccacgcc aaacccggtc agcatcagct gcacctccga cctgggcgcc aacgccatgt 1800
tcatgtcctt catcatcttc cacaccatcg ccgtgttcgt gttgcctggc attggtgtgc 1860
tactcaacca ctacggcgtc cgacgcaagc tctgcgccct gtcgctaacg gctcgagctg 1920
cccatggaga gctgccgctg cccattccca tcctgcgacg ccagacccac atggtcattg 1980
tgacgggctg ccccaacgcc cagcaagcgg cctgtggtgg gggcaccacg gcggacgaca 2040
cctcgaacgg gaacggcacg ggcaccggtg gcggtcccat ggccgtcagt ccgggcgaca 2100
tacagcttca cacactgcag ccgcgccagc caggctctgc cggatcggcc ctcgaaccgg 2160
gctcctaccg ctctagcaat cccatatcac ccaggtaagt cataaatatt gtgaagttat 2220
cagcaaaatc gttgtaaaca ggatgtatgc ttgatgtagt caaaagctat tagtttatgt 2280
taatcatacg ccccgttgac tttggtcaga atatatttca tcgaatggca gtacgtcatt 2340
tggtataagt taactaactt gcaaacccac ttagggccat gagggagatt cgcgcccact 2400
cgcagcgcca gcgcatcaac cgagcaggca ggggaccagc gacacccggc atcccactgc 2460
cccaaacctc cacgctgcgg tcgcgaaggc acctggctaa catgctcatc gcctcggccg 2520
tgatcttcat tgcctgctgg gcaccgcacg tcttctgcat cttctacaag aacttcggca 2580
```

acaatcagca gtgcagccag acctctgttt acttcagcct tctgctaggt aagtctgaat 2640
tggcaatcat tcccaatccg aatccgagtc cgaattggaa ctgactagtt gatcgcgttg 2700
caggctactt ctactcggcc atcagtccgg tgatctactg ggcgctcaac cacaattcgc 2760
ttcgacaatc tccctgcgcg cccatcatcc gcctgcgctc catgcagaac ttcctgaggt 2820
cgcgcttccg gacgcacacc gccccgccgc ctccctcgtc caccaacgag gcggcactgg 2880
gggcattcaa ccccaagctg atcaagctaa caccgaaaca gtatagagct caggcctctt 2940
cgcattatct ctactagtgt gatactaagc taaacttgat atacatatat acacatttat 3000
atacacgcat atatagtata ttttttacct acgaccccac cccccccctg acccgtcgat 3060
gccctcttag taccctctag tatcgaactc cgaacccgaa ctttgaccca accgaacgaa 3120
actatctaga tgtttttcaa aagctctacg caacaattgc ataatacgaa agatacatgc 3180
tacatataca catatataga gtacatacta ccttgcatag cagtgttagt cgacgcttat 3240
ttacacgatg ttgtgcgcct gaagctatgc aataaggaca ccggagtcct gccgcaactc 3300
ttgcctataa cgaggatagc actggtaacg aaagcaagaa aaaaaaaagg aaaaaaaata 3360
caaaaataac gaaacccttc tagtgacttg ttcgaattgt tgtttagtct tttaagctga 3420
aatattgtac tatgttcctt ctcatagata agcaataagt gaacgcttta agccaagtta 3480
aacaaacgga aaacgaatca aactaaacta aaactcgatc taaaactttt aaacgattta 3540
aacgtaaact cgaaagcata cctaacgaaa cattaataat tatatacata cgcgactaca 3600
tacaaataac aactaacaag tgaaaactaa aaacaaaaaa tgaatgtaaa tgtgagcaaa 3660
aaaacaaaag gataacattt agtattgtat acatattctt cgacttaatt ctttatacac 3720
tttttctaac catttttaat atcgttcagt aaactgaaca cggtgtatag ggcaattcaa 3780
gctgctggat atattcgcgt ttcagatgcc agagccaccc aaaggatcgg actctggatc 3840
gggtttcatt tcactccatt cagtggcgac cgtttcctaa tttgccaaac ttttgtacat 3900
gctccgctgc gagcgactct ctccttactc cacctggttt ttcgttcaat ccttgga 3957

<210> 170
<211> 1638
<212> DNA
<213> Drosophila

<400> 170
atgaacaggg acaatctgca gcagtggtgg gagaactcgt atcggcgcca acatccagag 60
cccaccgatg acctcggcct ggactccgcg gagctgcatc tcgccctcca ggagcccaac 120
cagctgcccg cggactacga ctatggcaac tttagtctag gcaatccgta tgacgtggac 180
tcagagcact ccatttcccc gctgacgctg ctcctgctcg ctgttagcta tggcctggtg 240
gttttcggcg gcgtcgtggg caactccacg ctcgttctga ccctctgctc cgcctcttcg 300
gtgcgtttgc gtaacccgct gctgctggcc gtctgcattg ccgatctcct ggtcaccggc 360
atctccgctc cggtcacgct cctcaatctc gcgatgaacc ggaggacgcg atcgctgccc 420
cttgtgctgt gcaaggtcat acactacgta caggtcatgc ccgtttcggc cagcactatt 480
tcgtttttca tgctctccct tgaccgctac gccacagtaa agcatccgcg attggcgcag 540
cttcgccagc ggcgctactt gcacgtctcc ctggcgctgc tatcctggct cgcctctgcc 600
gccatcagca ctccctttct gtttgcctac aaaatcatcg ccaagtcgat ggtggtcaag 660
ggcggtgggg cggcgaacac cacgccaaac ccggtcagca tcagctgcac ctccgacctg 720
ggcgccaacg ccatgttcat gtccttcatc atcttccaca ccatcgccgt gttcgtgttg 780
cctggcattg gtgtgctact caaccactac ggcgtccgac gcaagctctg cgccctgtcg 840
ctaacggctc gagctgccca tggagagctg ccgctgccca ttcccatcct gcgacgccag 900
acccacatgg tcattgtgac gggctgcccc aacgcccagc aagcggcctg tggtgggggc 960
accacggcgg acgacacctc gaacgggaac ggcacgggca ccggtggcgg tcccatggcc 1020
gtcagtccgg gcgacataca gcttcacaca ctgcagccgc gccagccagg ctctgccgga 1080
tcggccctcg aaccgggctc ctaccgctct agcaatccca tatcacccag ggccatgagg 1140
gagattcgcg cccactcgca gcgccagcgc atcaaccgag caggcagggg accagcgaca 1200
cccggcatcc cactgcccca aacctccacg ctgcggtcgc gaaggcacct ggctaacatg 1260
ctcatcgcct cggccgtgat cttcattgcc tgctgggcac cgcacgtctt ctgcatcttc 1320
tacaagaact tcggcaacaa tcagcagtgc agccagacct ctgtttactt cagccttctg 1380
ctaggctact tctactcggc catcagtccg gtgatctact gggcgctcaa ccacaattcg 1440
cttcgacaat ctccctgcgc gcccatcatc cgcctgcgct ccatgcagaa cttcctgagg 1500
tcgcgcttcc ggacgcacac cgccccgccg cctccctcgt ccaccaacga ggcggcactg 1560
ggggcattca accccaagct gatcaagcta acaccgaaac agtatagagc tcaggcctct 1620
tcgcattatc tctactag 1638

<210> 171

<211> 545
<212> PRT
<213> Drosophila

<400> 171
Met Asn Arg Asp Asn Leu Gln Gln Trp Trp Glu Asn Ser Tyr Arg Arg
1   5   10   15
Gln His Pro Glu Pro Thr Asp Asp Leu Gly Leu Asp Ser Ala Glu Leu
            20   25   30
His Leu Ala Leu Gln Glu Pro Asn Gln Leu Pro Ala Asp Tyr Asp Tyr
        35   40   45
Gly Asn Phe Ser Leu Gly Asn Pro Tyr Asp Val Asp Ser Glu His Ser
    50   55   60
Ile Ser Pro Leu Thr Leu Leu Leu Leu Ala Val Ser Tyr Gly Leu Val
65   70   75   80
Val Phe Gly Gly Val Val Gly Asn Ser Thr Leu Val Leu Thr Leu Cys
                85   90   95
Ser Ala Ser Ser Val Arg Leu Arg Asn Pro Leu Leu Leu Ala Val Cys
            100   105   110
Ile Ala Asp Leu Leu Val Thr Gly Ile Ser Ala Pro Val Thr Leu Leu
        115   120   125
Asn Leu Ala Met Asn Arg Arg Thr Arg Ser Leu Pro Leu Val Leu Cys
    130   135   140
Lys Val Ile His Tyr Val Gln Val Met Pro Val Ser Ala Ser Thr Ile
145   150   155   160
Ser Phe Phe Met Leu Ser Leu Asp Arg Tyr Ala Thr Val Lys His Pro
                165   170   175
Arg Leu Ala Gln Leu Arg Gln Arg Arg Tyr Leu His Val Ser Leu Ala
            180   185   190
Leu Leu Ser Trp Leu Ala Ser Ala Ala Ile Ser Thr Pro Phe Leu Phe
        195   200   205
Ala Tyr Lys Ile Ile Ala Lys Ser Met Val Val Lys Gly Gly Gly Ala
    210   215   220
Ala Asn Thr Thr Pro Asn Pro Val Ser Ile Ser Cys Thr Ser Asp Leu
225   230   235   240
Gly Ala Asn Ala Met Phe Met Ser Phe Ile Ile Phe His Thr Ile Ala
                245   250   255
Val Phe Val Leu Pro Gly Ile Gly Val Leu Leu Asn His Tyr Gly Val
            260   265   270
Arg Arg Lys Leu Cys Ala Leu Ser Leu Thr Ala Arg Ala Ala His Gly
        275   280   285
Glu Leu Pro Leu Pro Ile Pro Ile Leu Arg Arg Gln Thr His Met Val
    290   295   300
Ile Val Thr Gly Cys Pro Asn Ala Gln Gln Ala Ala Cys Gly Gly Gly
305   310   315   320
Thr Thr Ala Asp Asp Thr Ser Asn Gly Asn Gly Thr Gly Thr Gly Gly
                325   330   335
Gly Pro Met Ala Val Ser Pro Gly Asp Ile Gln Leu His Thr Leu Gln
            340   345   350
Pro Arg Gln Pro Gly Ser Ala Gly Ser Ala Leu Glu Pro Gly Ser Tyr
        355   360   365
Arg Ser Ser Asn Pro Ile Ser Pro Arg Ala Met Arg Glu Ile Arg Ala
    370   375   380
His Ser Gln Arg Gln Arg Ile Asn Arg Ala Gly Arg Gly Pro Ala Thr
385   390   395   400
Pro Gly Ile Pro Leu Pro Gln Thr Ser Thr Leu Arg Ser Arg Arg His
                405   410   415
Leu Ala Asn Met Leu Ile Ala Ser Ala Val Ile Phe Ile Ala Cys Trp
            420   425   430
Ala Pro His Val Phe Cys Ile Phe Tyr Lys Asn Phe Gly Asn Asn Gln

435 440 445
Gln Cys Ser Gln Thr Ser Val Tyr Phe Ser Leu Leu Leu Gly Tyr Phe
450 455 460
Tyr Ser Ala Ile Ser Pro Val Ile Tyr Trp Ala Leu Asn His Asn Ser
465 470 475 480
Leu Arg Gln Ser Pro Cys Ala Pro Ile Ile Arg Leu Arg Ser Met Gln
485 490 495
Asn Phe Leu Arg Ser Arg Phe Arg Thr His Thr Ala Pro Pro Pro Pro
500 505 510
Ser Ser Thr Asn Glu Ala Ala Leu Gly Ala Phe Asn Pro Lys Leu Ile
515 520 525
Lys Leu Thr Pro Lys Gln Tyr Arg Ala Gln Ala Ser Ser His Tyr Leu
530 535 540
Tyr
545

<210> 172
<211> 7112
<212> DNA
<213> Drosophila

<400> 172
gaatgataac taattcaata aaactacaga tttcaataac taaagcatac atgtatatgt 60
atctatgtat ctatgactat gtatctcgaa tcctccgctg aaactctaaa ctattcccat 120
gcagtttctt ttccttcgta tcgagtttca gagcacggct tgtagtgttg cacgtttttc 180
gtatttttg gcagctctgg ctggtttttt tcattttcct ttttttttg tactcccatt 240
tcctgtcacc ctgcccggtg taattaaaaa agtttttctc gacatttgct tgccatgtga 300
ggctccagga acaaatgcgt tacattgtgg accatccgat tgtaaaataa attttaagtg 360
caatacgaag tgcaaactca gccgtaacaa taagattttt gcaataatat ttaactttta 420
attgtaactt tggtaaagtc aatgtttcct ttttccaagt tggccaataa aaagagcata 480
aagcaaagcg aaaactttag tgcgtacttt tattacaaca ttaacttcat gctatatatt 540
ttatttaaat attttggctt ttctatccat tgtttttcag ttaatatctt aaacaaacaa 600
tttttgatgc actacaaatg ctgttgattt taataaacga tttatagctt acatggatta 660
ttagctatgt acatataatt acgaatattt ttaataaagg gcatacttaa taaattgaaa 720
ttctggaaaa cactcattta tttaacacta gcaaattagc gcacaaataa tgtgcaaagt 780
tcgttaatct ttaaagtgta tttgtgacca cagcacaatt aaacggtggt ccgaaattct 840
atatcttcta attccgcatt gttttttttt ttttgtgaat ttttcaggta acctccttgt 900
tattttggtg gtcacattat cacgacgact gcgctcaata accaactttt tcctcgccaa 960
cttggcgttt gccgacttct gtgttggact cttttgtgtg atgcagaatc tgtccatcta 1020
tctgatagaa aggtgagtga caattggcaa aaacaacatt atttcagtat taaattaacg 1080
tttcatcatg cacgagcgga gctgcaaaat tctcggaaag catatacgaa gaatacgtga 1140
actaacttca ctgtaaagtt gccaatagtt tttaatgtcc gttttcatac gcaaaaaatt 1200
taagtgaagc atttaagtga agcatttata gaaaagtttt tagttatagt aggttaaaaa 1260
ttaacacaaa agttaaagta gcgtttcagc tagtacagat agtagaatgt tatgtggtcg 1320
attaaattta tttatttatc agctgcagaa tgtggctgaa acggtatctg agtgcatacg 1380
tatgtatgca tgtgcctcga aaacaaagcc ccaaaaaaaa gtttcaactt tcaaccgtga 1440
aacatttacg ccggctgtca cataaaggaa actaaaactc cgcaggcgtc actttcactt 1500
caggcttttc aggcttttat ccccacgctc cgctttcact ccaatcatgt ttatttgcat 1560
tgtgtttcct taactttccc aaaaacacgt cgcagtgacg tgttcactta aaataaagcg 1620
tggagtgctg cttgaaaatt caaacaaaaa gaagcctttt tcagatttgc agggcaaatt 1680
aaaaaagtat tttgacaaat gagtttggct tacagtttct ataagttatt taatttaaaa 1740
ttttagtcgc agtgtcagaa ataaataatg cttcattaat ttaaatttgt gaaatgcatg 1800
tgtttaatat gtgtttaatt aaaagctatg ttctatacat ttagaaaaaa aagtaattac 1860
ctatctcatt agaattcttg aaatgaccat ttatttatgc acttaataat aaacagacat 1920
aatgatttat ttgcattccc gatgttaaat cctgaaattg gtttgcatat tttccatttc 1980
attttctaat atgcttttttg tggccaaaga actaaattat tgagatgacc catctgtgtt 2040
tggcataatt ttagctaaat tatgcgcaca acttgacgcc caccagaaat catatgcaaa 2100
atgagatgag attacgaaaa atcataattt atgcatttat aatttatatg tgcggcatac 2160
ggctgccagt gacaatttaa tttaatgatt ataaacatca agtcgtttta agaaagtttt 2220

```
tagttcatta tgttaaggaa tgatttgtcc cttttaaaga gagaaattga atcaccaggc 2280
aaatggggaa tctttttaag aaacgcatta tccaaaatct tggcgacaac cttcaggatt 2340
attatgttta ttatagtatt attgggcaca tttaaacgca tgagtgcgaa attacgtggc 2400
attgctaata acgaggcaca aatgcggttt tcgaatgtaa tttcagccga aatagtaata 2460
aaaaggaaaa agctttgcgc ctcataaaat actaaagaca ccttaatgtc gcattagttt 2520
tggcgggcgt gtcataacat tttatgtggc ttgtgctaaa tttggctctg ccaccaatat 2580
tgccaagtat atgctcttta aagagcgaaa cttttgccca taaataatta tttatgtcct 2640
ttcagctggg tctttggcga gttcctctgc cgcatgtacc agtttgtcca ctccctgagc 2700
tacacggcct ccattttcat cctggttgtc atctgtatgg agcgctactt tgccatcgta 2760
caccccatta cctgcaaaca gatcctaact gcggctcgtc tgagggtaag ttctaaccat 2820
aaacctcata ttcacggcca gtaataggca agtatgcatt ttaaatgagg tctatgtgta 2880
tgtgccctgc ttttttcaat tcccattttg atgggtcaaa agcggaggaa atctgtttat 2940
gtattgtatt ttgagacatt ctcgcaactt taataacttc atttgaatac aatttgcaag 3000
ctaatgaaac ttattagacc aagcccggag actcatgcaa attaaccggc agaacagaac 3060
cgactcgaaa ggcttaataa gccgtgggca taactttgtg gccgggaaaa cttttccgca 3120
ctcgaagggt gaagtggagg gagcgcagcc aaagggttga accacatagc ccggttgttt 3180
catcacagcc agacaaaaaa tacgtttcgg taatgacgat tgcactaatt aaactacaag 3240
agacataaag acaacgaaat gggctatcca ttaacatgca aattcagtgt gttcattcgc 3300
aatttaacaa attgtttgcg tgtttggcag ccgagcagac aaaacaaaac gtcaatgcgt 3360
gttttttttt ttgcagatgg tcattgtcac agtttggatt acatcggcag tttactccac 3420
gcccaagttt gttttcagca agaccattaa gaacatccac acgcaggacg ggcaggagga 3480
ggagatatgt gtcctggacc gtgagatgtt caactccaag ttgctggaca tgattaactt 3540
tgtgcttctc tacgttatgc cgctcttggt gatgacggta agtgtaacaa atgatacgta 3600
gcttgaagtt aataattatg tttaaccaat ccttggatag gatgttcata aattcatttt 3660
caatccaaaa gtgtttccta gtcacgttac ctagcattga gcagactata attaataatt 3720
aaatgtattt ttttgaattg cttattaaag tgattaatta tatcaataac aactatctct 3780
ctaatgagat agaatataaa taattttggg tacataataa ataatctatg aaatgatttt 3840
tcttcgtctg ttaaaaaagt gattaatgga agtgtagatt tccctttatt tttttaacga 3900
ctccattttt caatttcttc tgagaaattt gataagaatt gagaagcttt ttgcataaaa 3960
tcgccccttg ttccgcaatg aaatttgaat aatgaattta acgaacccat ttgcattgca 4020
ggtgctgtac agcaaaatcg ccatcgcctt gtggcgcagt tcgcgcggcc tcacgccgca 4080
tgtggtgcag catcagcatc agcagccgca gcagccatcc tgccaggaca ttggcatggg 4140
tatgcacaac agcatgtacc atcaccatcc gcatcatcat caccaccacc accagcatca 4200
ccagctgcaa tcggcggctt cgtcggcagg agtggtggga gtgggtttgg gtggtggtgg 4260
tgggggcggt cctggtccgt cactcgcttc cggcggcagc agcaccacgt ccttgtcccg 4320
caaacagagc agtaaatatg agaagcgcgg cgtcagcatc acggagagcc aggtgagctg 4380
ccaggagaac actgcatcag aaaaatcgcc ggttgagcac ggaggagtag ccttaaaagc 4440
ctaacaatag caataaaccg aaagtggagt cgaacctttg ttccctaatt ggccagaacg 4500
gctttaactt cgattgagtc cgcacgtgct gtgtggtaaa taatgttttt tgaaaatcac 4560
atcgctatcg gagcactaat ttattagata caccatttgt cggcctggcc aagttggcag 4620
tcgtaaaaac tttacccgta gctgttactc ccgcactttc gccgggccaa gtggcacggt 4680
gctcggagca ggtgtcaatt cacattaagg ccatcgcagg accatgccac aagctcggcc 4740
accgatgtga cacctggagt gtggagtggt ggagcgaagt gaccaccggc acagggcgaa 4800
actgcaagtg cgaggccaaa gttatggcaa aagtgttttt ttttttccat tatcggtggc 4860
cggacagcag gttcaattgg tcgggtgcac catggcgtat ttcttatttc ttatttcgag 4920
ctcgacagca gctttgccac acaaactggc acacacatca ttgttcttga tcagcatcag 4980
cagctcactt tgtcaataaa agccggcgga taaaaatcga gaatttgtga cacgatgaca 5040
ttcggaaaag ttgtgcaact gccgcagatt gaaaattggt aaacaccaaa ggcaaggccc 5100
ttgaattttt taatattcct cgaagtatca aatgtcaggg gagcactggt tcgatttcgg 5160
tatgttttcc gttttaaaag tcacaaaaat ggaagtttgg taggaaaaac aatggcggga 5220
acacctttat aaaactttta cataacttta gtttacgctg ttccttaact tatgaccccg 5280
atatacacag ataacaacct aactatcttt ctgttgaact cttttctctc tctctacctt 5340
tcgtttgatc aataaaactc tgtgctctcc ccctctacta actccccaat tataacaaca 5400
accacacaca atcccacacc acaaattagc tcgataattg caaggtaggt tgtaatcgaa 5460
tatcaaaatc gatatgataa tgttcaacaa ctgttgcggc cagcgaccta atgccattac 5520
acctgttaat ccgcacaggt gtccctggag gccgatcgtc ccattgtcag tgcctgccgc 5580
aagacgagct tctaccacca cggccacgcc catcaccagc gggcgggcaa cgcctcagtg 5640
ggcggtggca gcggcggagc cggggccgga gccacccaca tgtcccactc gtccagcaac 5700
gtcctacgcg cccggcgcgg cgttgtccgc atgctgataa tattcgtgct gacattcgcc 5760
ctctgcaatc tgccgtatca tgcccgcaaa atgtggcagt actggtgagt agtcccttcg 5820
```

tccacccacc acccataatc acaatccaat gcgaatcgga atcccaatcc ctatccctat 5880
cccttaccgc aggccattca tcagcccaca cggactgagt gtaaacaata ttttatggta 5940
tattgcaccg gcaatcggcg ggaatttgca tagtcctccg ttggacggga aattaacatc 6000
tgcgagtggg cagataagct ctagatggct tcctgctgga atcgaggcta gtaaacaatt 6060
aagtcatcaa tggacagagc tggagctgca gccagctttg ccgctggcct aaatgatttg 6120
aaagcggtct tccgttttcc aatcccactc ccttcccaaa gacgccagtt tcgctttctt 6180
ctcgcctttc aaaagcggtc gccaagaaaa gtccgtttta gccagccctt tggctatcag 6240
ttccgcttcc ggtggagcac tttcgaatag gtggacggac cgcttagttt gctttgaatt 6300
cgggcaattt ggtttttccg gctggcgggt gtggtgggcg tataacttga taagtgagtg 6360
ggcggtggaa gtgtttttgt ttgctaaatc gaggaggagt taattgctcc ctgggtagaa 6420
ttaaaaacgt gcacttaagc tgctgactgc tgtaaattgg aaataaaata aaagtcacag 6480
aaatgtcctt taagcgaaag tcgaaagaca aacgattgat tgaagatggg aaaagtcatc 6540
aagtgtgaaa atctggagac attggcaatt ttggcacgta agagaaacgt ttttattttg 6600
aacttttcat acagattcta ctaaaccaaa cgtttttgag tcaattgtaa attgtaaagc 6660
attttacatg ccgatgtgca ataaatccaa tttaatgttc ataagttcat gtttttggtg 6720
aataactact gaaataaata tcgatctcta gttctcatta gtgcaaatat tcagtgtaat 6780
tttgtaagaa aatacttcgc tgcagatttg caacgttgca cttttggcaa acgcaacgaa 6840
aacgaaaatg tttttgaatt tttgttcgag cactcaattt gcacaattgt aaaacgaatg 6900
gcggacaaat aaatataaat aatatctggc caaatacaag agaaatggat ttccacatcg 6960
ccacttttcc cgcagtgcca taattgagat ttatgtctag cccggaatgg gcagcgagac 7020
ttgtttactt gtttgaactt tcagctttta tttatttgaa gaattggttg cccccaggga 7080
tattggtatt gttaggtgga attgaaagtg tt 7112

<210> 173
<211> 1437
<212> DNA
<213> Drosophila

<400> 173
atgcagaatc tgtccatcta tctgatagaa agctgggtct ttggcgagtt cctctgccgc 60
atgtaccagt ttgtccactc cctgagctac acggcctcca ttttcatcct ggttgtcatc 120
tgtatggagc gctactttgc catcgtacac cccattacct gcaaacagat cctaactgcg 180
gctcgtctga ggatggtcat tgtcacagtt tggattacat cggcagttta ctccacgccc 240
aagtttgttt tcagcaagac cattaagaac atccacacgc aggacgggca ggaggaggag 300
atatgtgtcc tggaccgtga gatgttcaac tccaagttgc tggacatgat taactttgtg 360
cttctctacg ttatgccgct cttggtgatg acggtgctgt acagcaaaat cgccatcgcc 420
ttgtggcgca gttcgcgcgg cctcacgccg catgtggtgc agcatcagca tcagcagccg 480
cagcagccat cctgccagga cattggcatg ggtatgcaca acagcatgta ccatcaccat 540
ccgcatcatc atcaccacca ccaccagcat caccagctgc aatcggcggc ttcgtcggca 600
ggagtggtgg gagtgggttt gggtggtggt ggtgggggcg gtcctggtcc gtcactcgct 660
tccggcggca gcagcaccac gtccttgtcc cgcaaacaga gcagtaaata tgagaagcgc 720
ggcgtcagca tcacggagag ccagttggca gtcgtaaaaa ctttacccgt agctgttact 780
cccgcacttt cgccgggcca agtggcacgg tgctcggagc aggtgtcaat tcacattaag 840
gccatcgcag gaccatgcca caagctcggc caccgatgtg acacctggag tgtggagtgg 900
tggagcgaag tgaccaccgg cacagggcga aactgcaagt gcgaggccaa agtgtccctg 960
gaggccgatc gtcccattgt cagtgcctgc cgcaagacga gcttctacca ccacggccac 1020
gcccatcacc agcgggcggg caacgcctca gtgggcggtg gcagcggcgg agccggggcc 1080
ggagccaccc acatgtccca ctcgtccagc aacgtcctac gcgcccggcg cggcgttgtc 1140
cgcatgctga taatattcgt gctgacattc gccctctgca atctgccgta tcatgcccgc 1200
aaaatgtggc agtactggcc attcatcagc ccacacggac tgagtgtaaa caatatttta 1260
tggtatattg caccggcaat cggcgggaat ttgcatagtc ctccgttgga cgggaaatta 1320
acatctgcga gtgggcagat aagctctaga tggcttcctg ctggaatcga ggctagtaaa 1380
caattaagtc atcaatggac agagctggag ctgcagccag ctttgccgct ggcctaa 1437

<210> 174
<211> 478
<212> PRT
<213> Drosophila

<400> 174

```
Met Gln Asn Leu Ser Ile Tyr Leu Ile Glu Ser Trp Val Phe Gly Glu
 1               5                  10                  15
Phe Leu Cys Arg Met Tyr Gln Phe Val His Ser Leu Ser Tyr Thr Ala
            20                  25                  30
Ser Ile Phe Ile Leu Val Val Ile Cys Met Glu Arg Tyr Phe Ala Ile
        35                  40                  45
Val His Pro Ile Thr Cys Lys Gln Ile Leu Thr Ala Ala Arg Leu Arg
    50                  55                  60
Met Val Ile Val Thr Val Trp Ile Thr Ser Ala Val Tyr Ser Thr Pro
65                  70                  75                  80
Lys Phe Val Phe Ser Lys Thr Ile Lys Asn Ile His Thr Gln Asp Gly
                85                  90                  95
Gln Glu Glu Glu Ile Cys Val Leu Asp Arg Glu Met Phe Asn Ser Lys
            100                 105                 110
Leu Leu Asp Met Ile Asn Phe Val Leu Leu Tyr Val Met Pro Leu Leu
        115                 120                 125
Val Met Thr Val Leu Tyr Ser Lys Ile Ala Ile Ala Leu Trp Arg Ser
    130                 135                 140
Ser Arg Gly Leu Thr Pro His Val Val Gln His Gln His Gln Gln Pro
145                 150                 155                 160
Gln Gln Pro Ser Cys Gln Asp Ile Gly Met Gly Met His Asn Ser Met
                165                 170                 175
Tyr His His His Pro His His His His His His His Gln His His Gln
            180                 185                 190
Leu Gln Ser Ala Ala Ser Ser Ala Gly Val Val Gly Val Gly Leu Gly
        195                 200                 205
Gly Gly Gly Gly Gly Gly Pro Gly Pro Ser Leu Ala Ser Gly Gly Ser
    210                 215                 220
Ser Thr Thr Ser Leu Ser Arg Lys Gln Ser Ser Lys Tyr Glu Lys Arg
225                 230                 235                 240
Gly Val Ser Ile Thr Glu Ser Gln Leu Ala Val Val Lys Thr Leu Pro
                245                 250                 255
Val Ala Val Thr Pro Ala Leu Ser Pro Gly Gln Val Ala Arg Cys Ser
            260                 265                 270
Glu Gln Val Ser Ile His Ile Lys Ala Ile Ala Gly Pro Cys His Lys
        275                 280                 285
Leu Gly His Arg Cys Asp Thr Trp Ser Val Glu Trp Trp Ser Glu Val
    290                 295                 300
Thr Thr Gly Thr Gly Arg Asn Cys Lys Cys Glu Ala Lys Val Ser Leu
305                 310                 315                 320
Glu Ala Asp Arg Pro Ile Val Ser Ala Cys Arg Lys Thr Ser Phe Tyr
                325                 330                 335
His His Gly His Ala His His Gln Arg Ala Gly Asn Ala Ser Val Gly
            340                 345                 350
Gly Gly Ser Gly Gly Ala Gly Ala Gly Ala Thr His Met Ser His Ser
        355                 360                 365
Ser Ser Asn Val Leu Arg Ala Arg Arg Gly Val Val Arg Met Leu Ile
    370                 375                 380
Ile Phe Val Leu Thr Phe Ala Leu Cys Asn Leu Pro Tyr His Ala Arg
385                 390                 395                 400
Lys Met Trp Gln Tyr Trp Pro Phe Ile Ser Pro His Gly Leu Ser Val
                405                 410                 415
Asn Asn Ile Leu Trp Tyr Ile Ala Pro Ala Ile Gly Gly Asn Leu His
            420                 425                 430
Ser Pro Pro Leu Asp Gly Lys Leu Thr Ser Ala Ser Gly Gln Ile Ser
        435                 440                 445
Ser Arg Trp Leu Pro Ala Gly Ile Glu Ala Ser Lys Gln Leu Ser His
    450                 455                 460
```

Gln Trp Thr Glu Leu Glu Leu Gln Pro Ala Leu Pro Leu Ala
465 470 475

<210> 175
<211> 17847
<212> DNA
<213> Drosophila

<400> 175
ccttcgactt gagcggacag ttgaggttct cgaactaatt gcaaaattgt caatgattta 60
ttcagttggc aattaaggtg ccaactaacg gtgtggctgg cacggctata tcgatttgaa 120
aatattcaat aatcccggag ccttgataag ctcatcaaac atgcacatgg ccatcagcag 180
gggggggcaa gtttgtattt agagccaagt gtgtttggtg aagggggggtt agctgtatat 240
ttcagggttt ccccttttttt tgtcgagcgg tttgggggggg ggttttgcga ggggacacaa 300
aagtacaagt tgctatggat caagctcaag gcgtggcagt tgggcattag ctctatggtg 360
gagcacggca accaagcgaa caaaggtcaa ctcgacaagt gagttaaatg acaaggacct 420
caccttgcct tccttgggct tcggctttgg ctttggcccg tgcgtgtttg catgtttggc 480
tcgatgaggc gacgcgaccc caaaaatgaa actcaacttt ggctttgact gcggctcggg 540
attgtggcgt gcaagcatta tggcctttcg ctttcaaaac aatgccccca tcctgtttcc 600
cggttggcct agtggcactt gctccttgtt gttggttttt cgcattgccc ttggccaatt 660
gcgagtgaat tttgatgttt tgactggcca ttcctggggc catttgacaa ctctataaaa 720
ccggtgtcct taagcccgcc ctcgtccttg cgcccaaccc ccaattaccg tccaactacc 780
acccacctac accccgctac acccacccac aaagtaaacg taattgcccg gagagtttcg 840
cgactttcga cttggccagc aaactgaaaa agcgaaactt caaaaaaccc gacttgactt 900
ggcaactgga aaagtggaaa agcggaagca cgagctttcc caccgaaaag ctctcgtgct 960
ctaagctttc gtcctggcga ggcgaaaacg taaatgtcag tcatggagtt ccagtcgtcg 1020
gatacatgtg gaacacaacg aacacagcag gagcctctcg cgtctcatcc cccgaataca 1080
atgcgaaaaa gtctgtacaa aaaatagaca aataaggagc aggagcgaaa acaccactct 1140
gtgcataaga tataactccc gcagaactga atacctgcaa aaagaaatac gcgttaaatg 1200
gaattacggt gcgggaattg ataaagtgaa gtgtgttaca gctagaaagc aagtgaaata 1260
aaacgaatac attaataatt catgcccctc atagtcagaa gcttactagt ttgataagga 1320
attcaacgta aatctgatat agtaatgtag aaaagttgta caaagtaata tacatagcat 1380
atagcatata gcacagttac acagaacaca ccgcaaactc aaacaaagtt gacaaaaaaa 1440
aagctcacgc ccctcacatc tcagtttttt ctcgaattat ttgtgaaagt gcaacaaaca 1500
tatataaaaa ataaaaaaga aaagggagga gagagcacag caaaaaaagg cagtgacagc 1560
aggacctgca aaggaattgc catcatttgg ggattacgga attatcggga tccgttctca 1620
ttaaaatgga tttaaggcaa tgccgatgtg ctgagctgca tttgtgcatg ggattatcga 1680
ttgtaatgcg cagtcatcaa attatgtttg aacaacgccg ccgcgacagc gataatagca 1740
acaataacaa caacaacaac aaccaaatca acaacaaaca ctgatttggc ccctgacccc 1800
ctttggcagc accccctcca tccaacaccc cctacgaccc tgcccaagca aaaggagcag 1860
cgaaaaagaa gaagcgtctt atgaaacaat tgaagcaccg gccaaaggag caaaggagca 1920
gacaaagaca gagacaaaaa catcggcaaa ggataaaacg gaggccgaga agaagaggaa 1980
gcagtagtag caaagcagga aacaacaaca acaacaaccg caacatcgca tttttaaacg 2040
gcaatacgca aagtaagtga agtgtaaagt gagcgagatg gcaatagaaa ggtgtagagt 2100
gccgcgaaga gaggcggaga gaatacactt gcaaaaatgt gagcagctaa ttcccaggga 2160
tgttaaaaaa aattaaaccc tcatggttca aagattctag gagatggaca agtgatgaat 2220
ggtcagtgga aatatatgta tgttttaagc caagtgaaat caaaatgatt tgagaagtac 2280
aaataaaaga aagaaaggaa agaatgatat atgatattta ttggttgtag tacaggaaaa 2340
taatagttaa aatgatatcc cggattgatg gtcattacaa tgatcgtacc acaaaactct 2400
ttctaactca aattaaattg ttaaaattat ttttcaatac aaatttgata tatttttgtt 2460
ttttcgccga caagatgcat agatctttca tagaatctga aatctatatc ttttttatat 2520
tgcttttgag aagagtaaac cgatcatcat ctacgatcac tcctttttgt tctgtgtagg 2580
tatgaagaga gcaggtggga aggcagtgaa gtgaagtgaa gcatttgtcg catttgtgca 2640
gaattcccga cgtaattaat gtttaaagtc tgtgtgtgtg tatgtgtgcg aaggtttgta 2700
aactaggtca ctaggaagtt agccccttga cgaccatccc ctctcatcag cgcctcttca 2760
gcggcaactc tacgtcaaag gtgtcaactc gccaaatctg caaccacacc ttagttaagt 2820
gagccacccc gtagccacca aaattatcca cagcaccacc accttttatt cgtccttaac 2880
acaccgccat gattaacgcg cctcagtgcg attaactggc aaacaaattt ggagtgacgc 2940
tgcggcgatg tacatacata catatatgta tgcatgtttg tgctaaaccg acagcatttc 3000

cattcaataa cacacgcggc tcgtatgcaa acactgtaaa aaggttttat gctaaaaaag 3060
acgccgtatt tgtgcatgaa aaactgtcat tcatcatcca accaacacaa ataaccggta 3120
atggcatcgc cacattatgg ctttgtattt tttacatcct caaacacaaa aaaactttag 3180
ttgatttctg acttgcatcg cattcatcac ccactaaagt cgagcaatca atcctcaaaa 3240
tacgtatacg cactgttagt tcagttgatt aataggctaa ttatgaagga aactgaaagg 3300
aaacactaaa tcaataaatc ctggcaaatt agcatcagta tatatttgtg ttagtgcaaa 3360
catttaagtt ttctatcaat tttaaataca aatatgttca gaattaagat atatttcaaa 3420
gggaactatg ctttgtttaa ttaagattta aggttttata actataaaag ccttcaaaat 3480
gtgccatcta ttttatattc acttacaagc agatttcaat caagaccact ctttaatatg 3540
atatattatg tctagaacaa ccgagaaggc actcttaaat taaatatttt agtaccgtaa 3600
taaattatga aacgtgctga cggtaaaata tataatacaa ttcatacttt aagttatatt 3660
ttttttttgc cctatatggt tttacgtatc ggtgtgtatg tgagtacctc gcacacgcag 3720
gacatagaga taaaaaggct tctcccacat aaaagttcta ttctgttttt tttttttat 3780
agaaaaatat tgctggcagg agtttagaac attgccaaag gggatgcgga cagatgaagg 3840
ggcgggggcg tggcgggtaa aagccataaa actcaggcgc cgaagcgcaa aatgcccttg 3900
ccctgaacaa aaagaagaaa ttgaacagca acaaccaaca acaaaaacaa agaagaacaa 3960
gagcaaataa aatacaaact gcttatttat ttatttgttt cgacaaatct caattgaaac 4020
aaagcttaaa agcttcgaaa tcaaaactta aatagaaacc acacagacat ccatccagac 4080
gaataaaacg aaaacgatag agcgcgagac aaagagacaa ttactattaa atccaagttc 4140
aattcggtcc gcaaacaatt gaatttgttg cagatttgat ttggcttcat gtggattgcg 4200
ataactattt tcctgggatt agtatgggat tggattcatt ggcccgaccg cagtttgctt 4260
cctaatgaga tttgcttaat taaaatgaaa ttcaataact ggaattccaa tgaatttaaa 4320
tgttagagaa aaactctaat tcggactttt gtggctagcc attatttgca gcaaggggga 4380
cataaaaagc atttggcttt cgtttggatt agtcagcaag gtcaaatctc gaaactgacc 4440
taagctaatg caaaatgact tttcctgtca atcacaatct atatttaatt caagcgaaat 4500
tctctgccac atttcttgcg ctataaatcg cagtacccat tatagcatat caacaaccaa 4560
attaagacga aagtacaata aaatgcaact aatttcagtt ttaataatgg atttgctggc 4620
ggcttatttt tgtttcaaat ttggttattt tctttagaaa tgtgcgcata cattaagacg 4680
aacacaaaaa ccagctttgg ctcaatctct taaaccttta aaacgttata tgtacgcaca 4740
tatttatgta cgaacaaatg aaggaaaatc tcaagcctaa ttccgcatta agctgcaagg 4800
aaacacataa cccaaaagta tgcaacataa ataaaattta aatcagcacg tttatgaagt 4860
ggattatttg aactgaaatg cttattcaac gttattagat gagatttatg gccaggagag 4920
catttgaatt ggcagattat tcgggccaac caagtaatgg ccactatatc aaacaaatca 4980
gttatccttc aattgcaaaa tacacagtgc agacaaatag tttaaactgt cttagtggta 5040
atcaaaatca actttaatta tcaactaatg tatgtgtatt ttacgaaaca aaagtgcatt 5100
atgcatgaac gaatgatagc aatatttttc aatataaaag cttacgattg ttttcctttt 5160
gttatcaatt atctatcaca ctccacagac atactcctag atcataataa taagttatat 5220
aaatatagca tagattttga aaggttaaaa ggtttctcc ccaccactcc ctctaataat 5280
attgtatata ataaaatata tatacaaatg tatttcataa tttcattgcc atccaattcc 5340
gcatagtatc tggtataccc catctgcaga aagacaccga agtgtataaa ttgcttacgt 5400
ttgcccatgt agtcataaaa attgctttgg cacataaaac gaagcttgac gctcgctttt 5460
ggttttggtt ttggttttgg cctggtctag tctggtttgg tttcatttgg ccccttaccg 5520
ttcttgtttc ctttggtttc catcgctggt gggcttgttt ttttgcgaga gtcctggccg 5580
tgtccttctt atatttctat ttctatttgt atttgtattt gtatttgtat tccccgattt 5640
gtattttgtt cgctcgccaa ttcggtaatc acatttcaag cgttttattc gtcctgattt 5700
catacgattg ctgattcaag tttctgtttg tggtccttgc cttgcaggat atgacaagtg 5760
atggtgctgt tacgttttgg atatttttgc tttgtttaat cgcctcgccg cacctgcaag 5820
ggggcgtggc cgggaggccc gatgaactgc acatcggcgg catctttccg atagccggca 5880
aaggaggatg gcagggcggc caggcgtgta tgcctgccgc aagactggcg ttggatgatg 5940
tcaacaagca gccaaatctg ctgccgggct tcaagctcat cctgcacagc aacgacagtg 6000
aggtgagtag tcagtacact gaaaaaaaaa cgatatatac ctaaggatat tttgaatgta 6060
atgcgatgca atagtcctac agtccaagtt cgatacgaaa atagtttttt atgtttcctc 6120
ttaagctggc tttaataaga ataaacattt attgaagctt ttgaactttc ttataatttt 6180
cataggtaac taatatatac aatatattac attaaatacc caataagtat ccaataaact 6240
tcgaagcaac acagaacttg tttgcaacac gattaatatc actgcaaata atatcctgac 6300
caaaaacccc tcctgatgca ctgagttctg tttaagcctt ctgttcacga agcattttcg 6360
ctcagtgcac tagtaagttg gaaagcgata agccctactt atatgcagcc cctgctgcct 6420
cctctgctgc catgccacat acccttgggt gccacttatt ttgtttcgtt tttgtttctg 6480
ctcactgatg agtagtggag tgctggctga gcgctggcat tcattagcat atttctctag 6540
gggatcaggc tggggctggg gctggagctg ggcactgtta acaaaggaca aacagtttta 6600

ttatgccgca ttaaagggta ttgctgcgcg tgtgttccac aaattgaatt cgccgcacaa 6660
tccgccgcat gtccaacaat atttggcaat gcctcggctt aatgactgca cataaattgc 6720
gattgagatt cgactggagc acgttcgaaa cgttaccaac agcaattata ttggctgagt 6780
tcggttaatc gggcttttat gtgcgtgatt cgtaattgat aaagggtatt tagcattaag 6840
aaaaatgtgt gtatatctcg aacaaggtcg cccaataaat gtcgcccggt gttcattaaa 6900
tcataaatgt atttacattt attcgtgtag tgcaacgctg taagtttaat tatttcaatg 6960
cgaaaatctc tcatttccca gaaagtttct atgagatcaa taagtaaaag tttttgctat 7020
aaacttcgaa cttggttttt tgcaatcaaa agttttacgc attttatctt taatgccagc 7080
agcagcacga actgaacttc tgcacaaagc ttttcttgaa taaatcggta taatcaaagt 7140
gaacacagta ttgatcatat ttttttattt gatccttata taaaattgtg aaatttttgc 7200
attttgtagc taacttaaag agcgcggaat atgtttttag ctggcatatt tgttctaaat 7260
ggaaaaaata gagcatgggc acgcaataaa tcaacatttc aatagtatcc agttgtctgg 7320
atgctgtaac ttctgttggc ttttagagaa ccagctatta aacaaatacc cgccctccaa 7380
acggtcatgc atcataataa atggccagta accgatgtag tttccgtttt ccagttccca 7440
accaagttgc gcgctatcaa tgccccttgg cgcgtatctc taatacaata catataaaac 7500
caaaaccaaa actatcttta tacattttca tatttcgcac atatttttttg tttttttgtg 7560
aatttcgcat tttaagggtc tgcggtacag ttgttgcggt ttccactact tcggcggctg 7620
ttttgggcct gcctattttt aagcgcgaac ctcattagtt ttaagtgagc gaaagcttaa 7680
ctacttgttt tgagttttga atgaggtcac tttgccccgg ccgagccgta gctcatcatt 7740
tgagttattt ttattcaagg ctgccaggag ggacagtcac tactactaca tacatatatc 7800
ctacgaacat atacgaacat atacacgccc atttacgagc cagatacgga catatagata 7860
ggaacagagc cggagaggta aaccaagtcc tatccctggt catatctgtt cctatctggc 7920
gtggtctggt ctggtctggt ctggttagtt cagttcagtt cagttcgctt cgcttcgctg 7980
gactgagcct gagagccagg cgcgttatgt ttgtggctta ggcaaacagc ttccgttaaa 8040
tgctgcataa ttagttgcca tctccaggtc ccacagccac cttctcctcc ggcacatcca 8100
catccacact tactcccaca acgcaacgca ccaccccatc ccgcccacat ccacgtcact 8160
tggcaagcga caaggggcaa aggtcgcacg caatgtttgc ctttgccgct gccttttgct 8220
tgacctactt tgatgtgtga aatgtcccaa agaattctgc acaaaagcac tgagagttga 8280
acaaaaaaat gtataaacaa ggcaggttgg ttatcctgta ataagcgaaa tccctagtcg 8340
gcttaagtcc aagatcacac agaacaggtg aatacttttg tggcgaattt cgaatcgaac 8400
tttgattttt taatggccaa atataggtgc tgaaaaatcg tattattgca tttgcaaagc 8460
tgccactttc tatcggctaa gtgccttcat gcgaaccact gaaatggctt aaaccggttg 8520
tcaagcttta ggaataagcc ggaaattgat ttaacagcgt tgggataaaa atatgtgctc 8580
ctaaaaggtt tttgcttttt ctttttttaat taaagtttaa ttaaatcaca acaaagcatt 8640
caattattta aactcacatt ttcaattgca atacagaaaa ggtaaatata atggctaagg 8700
taatcaaaaa ttatataaat aggcatctac agaatctgcg agttttaatt taaattgaat 8760
attgttttgt acaatgtaat aatattccta aatgtgcttt gtattataga agaacattct 8820
actaataaat catttagcta ttttttggtt aatattaaaa gtactatagt actatgtact 8880
atcagtttta tgctgcttag ttgttctata aaatcagtat gcctctgccc acgaaaaggg 8940
aaactattaa attagagcgg acatcgtggc atgcctcggt acccaaagcc aaagcacaca 9000
cccagatctg aatagatgga gtagtacgaa caggataata cttatgcata aacagggact 9060
ggacatgaga aggaaatgaa aggaaaaaaa aaaccaatta gaatggttgt ttacacgcaa 9120
accaagtaag aaatatgaaa attgtcgtca gaatggggtt tgggccaaag taaaggaaac 9180
gccgaacagg gaaacgagtt tgccatgtaa atgaatgaat gcgagaatga ataactgaat 9240
gaatgaatca tttatgtggc caatgtggat ggatacacgt atgcccacgt atgtatgtat 9300
gtatattcac aacttcggca caagggaaag ggacacacct tcgtgaatta ttcattgtca 9360
tgaatgaatc ttgtttaatt gataaatgct atgacatggg gattaatgaa aagcataatg 9420
gaaaaatatc aaccaggtgg cacataaatt aggcaggact tggaatatgg aaatttaact 9480
agctacataa atattgctac aaattcgtat aataaatcac gctaaatgca acaattaaca 9540
tgcatagttt aaaatttgat tgattataat tttcaatttt ctctgccagc aagtttttaa 9600
ttaagcagct taatttcgta gtttgctatg tcaaatccag cattccatcc agttgattaa 9660
ttttcatttt caattatatt ggcgaaaatt ggacgagatg acgaaaattc cttctgggcc 9720
agttttgtgc attaaattca atttaaaagc catgccaatg tcataaacgc cccgcatgaa 9780
tcacaaagcc aattaatcat atcaataatt atcttccaat gaagacagac acataaataa 9840
caggaaattc gctattaata actcctgaag ggtcgttgct tagcataatc ataatcaata 9900
catcgcgttg gccatcaaac gaaacctaca tatccgccca tgtttatgaa tatttatcga 9960
gcaccgacgg atgagagtat aagtattctt ggcaacttgg ctaattccgc tttccccatc 10020
cccatcccac cattttcatt cacagtgtga gcccggtttg ggcgccagcg tgatgtacaa 10080
tctgctctat aataaaccgc aaaagctgat gctgttggca ggatgcagca cggtctgcac 10140
cactgtagcc gaggctgcca aaatgtggaa tctaattgtg gtaagtgaaa tttgcatagc 10200

```
gccagaatgg ctggaatata tacatacata cttcgaaata ctttacctac atgccactgc 10260
ccctccacag ctctgctacg gggcctcgag tccggctctt tcggatcgca aacgattccc 10320
cactctattc cgcacccatc catcggccac ggtgcacaat ccaacgcgca tcaagctgat 10380
gaagaaattc ggctggtccc gggtggccat tctgcagcag gcggaggagg tctttatatc 10440
ggtaggtaat cggtgctaac taattacatt tccatattaa cttcgaaaac ttttatgaac 10500
agaccgtaga ggatctcgag aatcgatgca tggaggctgg cgttgaaatc gtaactagac 10560
aatcatttct atccgatcca acagacgccg tgcgcaattt gcgacgccag gatgcacgca 10620
tcattgtggg actcttctat gtggtggccg ccaggagggt gctctgcgaa atgtacaaac 10680
agcagctata tggccgagct catgtgtggt tctttattgg taaatacaaa ttatactact 10740
tagataccga tatctcatat gcatagatac tttcacagag tcttcttgta actaaactac 10800
ttttttttta ggctggtacg aggacaactg gtacgaggtg aatctgaaag cagagggcat 10860
cacctgcact gttgaacaga tgcgaatagc tgccgaagga catctgacaa cggaagcgct 10920
catgtggaat cagaacaatc agacaactat atccggaatg actgcagagg aatttcggtt 10980
agtactacta ctctatccat tgctaaagtt tttgcttcaa tgtgccttct ttttgcgcag 11040
acatcgactg aatcaggcgc taatcgagga gggttacgac attaaccacg atcgctatcc 11100
ggagggatat caggaggcgc cactcgccta cgatgcagtg tggagtgtgg ctttggcttt 11160
caacaagacc atggaacgat tgacaaccgg gaagaaatct ctgagggatt ttacctatac 11220
ggacaaggag attgccgatg aaatctacgc tgccatgaac tccacacaat ttctgggtgt 11280
atcggtcagt ttgtgaaact aatttcataa tacttagtta gttacttatt tttttatttg 11340
cagggtgtgg tggcattcag ttctcagggc gatcgtattg ctcttacaca gatcgaacag 11400
atgatagacg gcaagtacga gaagttgggt tactacgata ctcagttgga taacctatcc 11460
tggttgaata ctgaacagtg gattggtggc aaggtaagca aatcagttag cgttgaaaaa 11520
acaatcatat aatcaaaatg tcggacgatg ataaaaatca aaagaagaca aagcatgtaa 11580
cgactgaacc gaataagaac ggaaatccgt ccaaagacga cgttaataca agaggatcat 11640
taacaagctt aagtgagaaa aatgatgaat ccaaaaatcc atcatcatct gcaagtcttc 11700
agcaagctcc taaggtagca ccaaaaccga aaaagttatc aatttcggat gctggtaaag 11760
acacagtaac ccaaaaagtc aaagaaaatg aagagccagg attttctaaa caatttgaaa 11820
aagaatctat tggtgtgcga gcaaatagaa acagcagtgc aacaaaaaag gaaaatgaaa 11880
agctattagt aaagaccgtg cccggaaaat ctcttataaa agaatctaac gatgaaaatg 11940
tggaaccatc tagacgcaca aaaagtcagc cagttggaaa aaaagtttac gaggaatctt 12000
caacgaggcg agttcgtgag ccatttgatt cttttgatag ggaaaagtac ctgtctgata 12060
tgatcgatta tgatcgatcg tccgaggacg tgccggaaaa aaaactatct tgggaaccag 12120
acagtacact tcgacgtcgt tttgtcagca ataatgaata tcattctagt cttgaaggtg 12180
aggaggaaga gatcgattta gattcagttg gatcgtcaac aatgcgaggt agctatcgaa 12240
tgccacagaa tgccgaggat gaacgacctg ttatcctagc tgaaataatc aatatgggaa 12300
agttgaagga gctcaaggag caacgcgaat caactaacgt atttcgtaac aatcgtagag 12360
atgataaaac cattgatgaa gtgaagggtt ccaagatcac aggaaccagc aaacaacatg 12420
ttgaagaagt taaatataat gaaacagcaa gtaatgaaga tgatgacaat gcacaggtaa 12480
aacatatagg tacaatcaca aaaatcttta gcattgcgca aagaatgaaa gagaagaaaa 12540
gctccaaaga aagtgaagaa gacgagaata aggacgagaa taaagatgta aaaaaagatg 12600
aaaataaaga cgatgaacag aaacttcccc cggaaccaga gatcgaattg gaggccaaaa 12660
aagcttggac ttttccgatc gcagagacat cacctgatat ccctggagaa aatcagatcc 12720
tgcaagaata ccaaaaggtt aaaattggca aaagaaacag gaaatacaga tgtataggca 12780
tcaatacgga tataagcaga aaatctaaga ttaaaacatt tatcagcaag tctgcaggtg 12840
atgacttagt agtataccat gatggtggaa gattgcgaga tataggacaa accactgaaa 12900
atcttaagaa tcaacgaaag aatcataaac acaaaaccga accagatgat atacccgttg 12960
atataggaca cagcgatgat cgagtaaggg aaattggagt gaatactaaa aagttaccaa 13020
aaatcataat accaccaatt gccgagatgc atgtgcacaa aaatggcaaa ttaagggaca 13080
ttggaaccag cacagataag ccattttggc caatcgatga cggaactgat gtaatctaca 13140
tgcatcccat taagactgat aggaagaaat taaacaaact gattgtggat ccaccgccgg 13200
ataatggacc ctacaagatg ccaaccaaag aagatagacg aacgtactac aaaggttgtg 13260
aatatcattt tcctggccga acggaatggc gcaggctttt cttcaacaaa atccatggaa 13320
aatacaagct acgccgtcca agtcattggc tctacacact agttttttcc gttttatata 13380
ttttgtttgt cattatcttt tccatggcct ggttcgattt catcaaagat gacgcttcca 13440
gaaaggtacc catgataaag atggcacaac catttattag tttcactccc atcggaccac 13500
gaactaatcc caaggcggta tcctttgatc ctagaaacag cacagaagtc atggaaaagt 13560
acgctggtat aatggctctt ctggaaaaat acggcgacta tggacacaat ccgcgttttg 13620
gaacgtgtac tgcgaatgaa aagtttggct atccaagcgg cgaaccgtgt gtttttctta 13680
aagttaatcg gattattgga tttaagaccg aaccctacat caattccgat gagcttgtca 13740
aagccaagat cgatgaggtt gaatttacgg ccttaaaaag gttattggaa aatacaacaa 13800
```

```
cagaggaggg tcatttgaat cgcacttgga taacatgccg ttcagataaa gataagaatg 13860
ttttgattga atttcatcca gagccggcga taagaacgga atacaccgat attgaagaga 13920
aaatcgagta tattgcgaac gagggtaaaa agtcattttt cggcccaaac gatgtgaatc 13980
gaatcgtggc tcttaagatc aaaaatctca aagcaaatga acgtgtccat ataaattgca 14040
aaatgtgggc tcagaatata catcatagga aagagggtta tggccaagtt tcgtttttttg 14100
tgttattagc tacaaatgaa aaccgcgaga gggtggaaaa attactccga accgactact 14160
cattataaaa tacataaata taactttcta ataaaatgtg ttatatgata tatgtatgta 14220
ttatttataa catctataat ccaggttcct caagatcgca caattgtcac ccatgttcta 14280
cgcaccgtgt ccttgccatt atttgtgtgc atgtgcacaa tatccagttg tggcatattc 14340
gttgccttcg ccttgatcat ctttaatata tggaataagc atagaaggta actatttaat 14400
tggtagtgct cgtgcagaac taactaaggg ttatataatt tctagagtaa tacaatcctc 14460
gcatcccgtt tgcaatacga tcatgttatt tggtgtcatc atctgtctaa tatctgtcat 14520
cttactgggc atcgacggac gctttgtcag ccccgaggaa tatccaaagg tgcgagttac 14580
ttaccttaaa agttctcttc atttttatta aagttaattt gatataaaat tatccattag 14640
atatgtcaag cgcgggcttg gttactatcc accggtttta cactagcata cggtgctatg 14700
ttcagcaagg tctggcgtgt gcatcgtttt acaacaaaag caaaaactga cccaaaggtg 14760
ggtttaaatt aaattaacta caactatata ttagttgcta atcattttgt ggaattgcag 14820
aaaaaagtgg aaccttggaa gctatacacc atggtttcgg ggctattatc aatagattta 14880
gtgatattac tctcatggca gatctttgat ccgctgcagc gttatctcga aacattccca 14940
ctcgaagatc cagtatctac tactgatgat attaaaatac gtccagagct tgagcattgt 15000
gaaagtcaac gcaactccat gtggttgggt atgttgtgat aacgatccat tatattgtat 15060
ataatatatt tatattattt tgcaggtctt gtatacggct tcaaggggct aatcctggtg 15120
tttggcctct ttttggcgta cgagacgcgc tccattaaag tgaaacagat caacgattcg 15180
cgttatgtgg gcatgagcat ctataacgtg gtcgtccttt gcctgataac agctccggtg 15240
ggcatggtca ttgcatcgca acaggacgcg tcctttgcct tcgttgctct agctgtgata 15300
ttctgttgtt tcctaagcat gctgctgata tttgtgccaa aggtcagcca ctacactgaa 15360
ctaaaagaag tagcacattt gtaatgatct ctattcgcac aggtcattga ggttatacgt 15420
catcccaagg ataaggccga atcgaaatac aatcccgatt cagccatatc gaaagaggac 15480
gaagaacgct atcagaaact tgttaccgaa aacgaggaat tgcaacgatt aataacacag 15540
gtatgctgaa ttatggtaga ttatcttaat aaaatcgtat tttatttaat tttgtttaaa 15600
gattaaattc aactgttcgt tgaacgttaa tctaaaagta tttcggtctt atcataattt 15660
gaacaaaagt taaggattgg aatttccagc ggccccagtg aattccattc agcatgccgg 15720
tttcaccttc tctaaaataa agtccattta gcgagctgca atatgaagta ttatcaacct 15780
cttataaatc aaattttaat ttgataatta atttaccttg acaaacatga gtgaaaccac 15840
caaccaccgc catgctcggc tgcacaattt ttactagatt catcattatc acgatcaaaa 15900
gttgagaacc ttttatttat atggtatctt agggagtccc cagctgttcc agaatatttt 15960
cccacccgtt ccagtttgta taattcagtt tcactgtcca cttgaaagtc atcaaagtga 16020
gcgtatactg tagcgccagg accatgcttc aattgaatga aaagttcgtg tggttgttct 16080
ctagtcatta ggtaaagctt ctgaagaccg atgaaaaatt cgccattcgg atcgccaaat 16140
ccatctttgt aactaaacca cgattggtta aaattaacac ttccatctag tcgtctttga 16200
atgacaatcc aatctcttgc cgtcgttttg cactgggtca cttgaaaagg ttcttcttct 16260
ggcaacatta gctggtgaat tccatttgga ctgccactag gacacgtgtg tggtgaactc 16320
tggcccactt cgaacagcaa agtccataaa aatagtacaa aaaagcacga tttcattgta 16380
ctcgacatgc gtactgttag aagatcggtg gcttactgaa gtatctacca tttaaaaggg 16440
gaacattggg acgcagtttt gtggttttgc tacctttgac cgcattgctt gatatgcata 16500
tattctattt aggtcggcat ttaagcaaat ggattccaaa aaagtgaatg gagactattt 16560
aaaagataga actacatttt tattgtttaa atgtttgttt gttttgctta tatttttatt 16620
gttttataaa cattgattta ttaattttcc gcttgcagaa ggaggaaaag attcgagtcc 16680
tgcgacagcg tctggtggag cggggcgacg ccaagggcac agaactgaat ggtgcaacag 16740
gtgtcgcctc cgccgccgtt gcaacaactt cgcagcccgc ttccctcatc aactcatcag 16800
cacatgccac gcccgcagcc acactcgcaa tcacacaagg tgagtagagt acctggtgca 16860
tggaactatg catcccaaat gccccttagc cagttgcatg tggtaatcga tactagttac 16920
gttggtttag cttgacgttc ttaatcgtta tgttctcgat tttaactaga aatcttagct 16980
ctattcgtcg gcattgacac aattgattag gaacaacaat ataaatggag ttcaaaggtt 17040
atttgcttat gctaacgcta tgtaaactcg attttcgcac aaaaaaccaa aggcaattta 17100
tcagtttctg gaatgtctat tgctttgtgt gcatacatat gtatcttaag cacttttgca 17160
gtttttgagc ttatttatat cttcagcact ttagcacttt tttaataaat atcactgctt 17220
ctgcaccctt ttttcacatt ttagcaagat tattcaaaac atactaatac gtaatatatt 17280
tatcttacag atactcatga gcatcgtaca tgaaaaaact tgtcaatttc ctttgcgtga 17340
taggtgaatt atacaaaaga tatgaaatgc aaataatgat gctaacggaa ggatggattt 17400
```

aaacaaacac gtacttaaag aactttgtat caactttaag ctgagctaaa taagtgaatt 17460
ggtagcagca attttgtaga ataatgaaat atgtattgta ccttaaacgt gtttagcatt 17520
aatattgaaa acgaagtaat tttaacttaa aagaaaacaa aacaaaaatt cgacactagc 17580
atctaaacta agcgaattgt taggctttaa agttgcaaca cagaaaccaa aaaaaattct 17640
agcaaatcaa gagaattcat tacgaaaaga tcgtgttaaa gttagtaatt tttaaatgca 17700
ggtagagaag acaaattgtt gattttgtgt ttagtaaata ctgttttttt agttaaacga 17760
attatccttt cgaacctcat ttgctcagcg ctgtctccta gagctctcta gttaattata 17820
tctgcgtaga tgaaaattcc ccaaaag 17847

<210> 176
<211> 5167
<212> DNA
<213> Drosophila

<400> 176
tcatggagtt ccagtcgtcg gatacatgtg gaacacaacg aacacagcag gagcctctcg 60
cgtctcatcc cccgaataca atgcgaaaaa aactgaatac ctgcaaaaag aaatacgcgt 120
taaatggaat tacgcacccc ctccatccaa cacccccctac gaccctgccc aagcaaaagg 180
agcagcgaaa aagaagaagc gtcttatgaa acaattgaag caccggccaa aggagcaaag 240
gagcagacaa agacagagac aaaaacatcg gcaaaggata aaacggaggc cgagaagaag 300
aggaagcagt agtagcaaag caggaaacaa caacaacaac aaccgcaaca tcgcattttt 360
aaacggcaat acgcaaaaaa aatattgctg gcaggagttt agaacattgc caaaggggat 420
gcggacagat gaaggggcgg gggcgtggcg ggatatgaca agtgatggtg ctgttacgtt 480
ttggatattt ttgctttgtt taatcgcctc gccgcacctg caaggggggcg tggccgggag 540
gcccgatgaa ctgcacatcg gcggcatctt tccgatagcc ggcaaaggag gatggcaggg 600
cggccaggcg tgtatgcctg ccgcaagact ggcgttggat gatgtcaaca agcagccaaa 660
tctgctgccg ggcttcaagc tcatcctgca cagcaacgac agtgagtgtg agcccggttt 720
gggcgccagc gtgatgtaca atctgctcta taataaaccg caaaagctga tgctgttggc 780
aggatgcagc acggtctgca ccactgtagc cgaggctgcc aaaatgtgga atctaattgt 840
gctctgctac ggggcctcga gtccggctct ttcggatcgc aaacgattcc ccactctatt 900
ccgcacccat ccatcggcca cggtgcacaa tccaacgcgc atcaagctga tgaagaaatt 960
cggctggtcc cgggtggcca ttctgcagca ggcggaggag gtctttatat cgaccgtaga 1020
ggatctcgag aatcgatgca tggaggctgg cgttgaaatc gtaactagac aatcatttct 1080
atccgatcca acagacgccg tgcgcaattt gcgacgccag gatgcacgca tcattgtggg 1140
actcttctat gtggtggccg ccaggagggt gctctgcgaa atgtacaaac agcagctata 1200
tggccgagct catgtgtggt tctttattgg ctggtacgag gacaactggt acgaggtgaa 1260
tctgaaagca gagggcatca cctgcactgt tgaacagatg cgaatagctg ccgaaggaca 1320
tctgacaacg gaagcgctca tgtggaatca gaacaatcag acaactatat ccggaatgac 1380
tgcagaggaa tttcggttag tactactact ctatccattg ctaaagtttt tgcttcaatg 1440
tgccttcttt ttgcgcagac atcgactgaa tcaggcgcta atcgaggagg gttacgacat 1500
taaccacgat cgctatccgg agggatatca ggaggcgcca ctcgcctacg atgcagtgtg 1560
gagtgtggct ttggctttca acaagaccat ggaacgattg acaaccggga agaaatctct 1620
gagggatttt acctatacgg acaaggagat tgccgatgaa atctacgctg ccatgaactc 1680
cacacaattt ctgggtgtat cgggtgtggt ggcattcagt tctcagggcg atcgtattgc 1740
tcttacacag atcgaacaga tgatagacgg caagtacgag aagttgggtt actacgatac 1800
tcagttggat aacctatcct ggttgaatac tgaacagtgg attggtggca agaagacaaa 1860
gcatgtaacg actgaaccga ataagaacgg aaatccgtcc aaagacgacg ttaatacaag 1920
aggatcatta acaagcttaa gtgagaaaaa tgatgaatcc aaaaatccat catcatctgc 1980
aagtcttcag caagctccta aggtagcacc aaaaccgaaa aagttatcaa tttcggatgc 2040
tggtaaagac acagtaaccc aaaaagtcaa agaaaatgaa gagccaggat tttctaaaca 2100
atttgaaaaa gaatctattg gtgtgcgagc aaatagaaac agcagtgcaa caaaaaagga 2160
aaatgaaaag ctattagtaa agaccgtgcc cggaaaatct cttataaaag aatctaacga 2220
tgaaaatgtg gaaccatcta gacgcacaaa aagtcagcca gttggaaaaa aagtttacga 2280
ggaatcttca acgaggcgag ttcgtgagcc atttgattct tttgataggg aaaagtacct 2340
gtctgatatg atcgattatg atcgatcgtc cgaggacgtg ccggaaaaaa aactatcttg 2400
ggaaccagac agtacacttc gacgtcgttt tgtcagcaat aatgaatatc attctagtct 2460
tgaaggtgag gaggaagaga tcgatttaga ttcagttgga tcgtcaacaa tgcgaggtag 2520
ctatcgaatg ccacagaatg ccgaggatga acgacctgtt atcctagctg aaataatcaa 2580

tatgggaaag ttgaaggagc tcaaggagca acgcgaatca actaacgtat ttcgtaacaa 2640
tcgtagagat gataaaacca ttgatgaagt gaagggttcc aagatcacag gaaccagcaa 2700
acaacatgtt gaagaagtta aatataatga aacagcaagt aatgaagatg atgacaatgc 2760
acaggtaaaa catataggta caatcacaaa aatctttagc attgcgcaaa gaatgaaaga 2820
gaagaaaagc tccaaagaaa gtgaagaaga cgagaataag gacgagaata aagatgtaaa 2880
aaaagatgaa aataaagacg atgaacagaa acttcccccg gaaccagaga tcgaattgga 2940
ggccaaaaaa gcttggactt ttccgatcgc agagacatca cctgatatcc ctggagaaaa 3000
tcagatcctg caagaatacc aaaaggttaa aattggcaaa agaaacagga aatacagatg 3060
tataggcatc aatacggata taagcagaaa atctaagatt aaaacattta tcagcaagtc 3120
tgcaggtgat gacttagtag tataccatga tggtggaaga ttgcgagata taggacaaac 3180
cactgaaaat cttaagaatc aacgaaagaa tcataaacac aaaaccgaac cagatgatat 3240
acccgttgat ataggacaca gcgatgatcg agtaagggaa attggagtga atactaaaaa 3300
gttaccaaaa atcataatac caccaattgc cgagatgcat gtgcacaaaa atggcaaatt 3360
aagggacatt ggaaccagca cagataagcc attttggcca atcgatgacg gaactgatgt 3420
aatctacatg catcccatta agactgatag gaagaaatta aacaaactga ttgtggatcc 3480
accgccggat aatggaccct acaagatgcc aaccaaagaa gatagacgaa cgtactacaa 3540
aggttgtgaa tatcattttc ctggccgaac ggaatggcgc aggcttttct tcaacaaaat 3600
ccatggaaaa tacaagctac gccgtccaag tcattggctc tacacactag tttttccgt 3660
tttatatatt ttgtttgtca ttatcttttc catggcctgg ttcgatttca tcaaagatga 3720
cgcttccaga aaggtaccca tgataaagat ggcacaacca tttattagtt tcactcccat 3780
cggaccacga actaatccca aggcggtatc ctttgatcct agaaacagca cagaagtcat 3840
ggaaaagtac gctggtataa tggctcttct ggaaaaatac ggcgactatg gacacaatcc 3900
gcgttttgga acgtgtactg cgaatgaaaa gtttggctat ccaagcggcg aaccgtgtgt 3960
ttttcttaaa gttaatcgga ttattggatt taagaccgaa ccctacatca attccgatga 4020
gcttgtcaaa gccaagatcg atgaggttga atttacggcc ttaaaaaggt tattggaaaa 4080
tacaacaaca gaggagggtc atttgaatcg cacttggata acatgccgtt cagataaaga 4140
taagaatgtt ttgattgaat ttcatccaga gccggcgata agaacggaat acaccgatat 4200
tgaagagaaa atcgagtata ttgcgaacga gggtaaaaag tcatttttcg gcccaaacga 4260
tgtgaatcga atcgtggctc ttaagatcaa aaatctcaaa gcaaatgaac gtgtccatat 4320
aaattgcaaa atagtaatac aatcctcgca tcccgtttgc aatacgatca tgttatttgg 4380
tgtcatcatc tgtctaatat ctgtcatctt actgggcatc gacggacgct ttgtcagccc 4440
cgaggaatat ccaaagatct ttgatccgct gcagcgttat ctcgaaacat tcccactcga 4500
agatccagta tctactactg atgatattaa aatacgtcca gagcttgagc attgtgaaag 4560
tcaacgcaac tccatgtggt tgggtcttgt atacggcttc aaggggctaa tcctggtgtt 4620
tggcctcttt ttggcgtacg agacgcgctc cattaaagtg aaacagatca acgattcgcg 4680
ttatgtgggc atgagcatct ataacgtggt cgtcctttgc ctgataacag ctccggtggg 4740
catggtcatt gcatcgcaac aggacgcgtc ctttgccttc gttgctctag ctgtgatatt 4800
ctgttgtttc ctaagcatgc tgctgatatt tgtgccaaag gtcattgagg ttatacgtca 4860
tcccaaggat aaggccgaat cgaaatacaa tcccgattca gccatatcga aagaggacga 4920
agaacgctat cagaaacttg ttaccgaaaa cgaggaattg caacgattaa taacacagaa 4980
ggaggaaaag attcgagtcc tgcgacagcg tctggtggag cggggcgacg ccaagggcac 5040
agaactgaat ggtgcaacag gtgtcgcctc cgccgccgtt gcaacaactt cgcagcccgc 5100
ttccctcatc aactcatcag cacatgccac gcccgcagcc acactcgcaa tcacacaagg 5160
tgagtag 5167

<210> 177
<211> 1713
<212> PRT
<213> Drosophila

<400> 177
Met Trp Asn Thr Thr Asn Thr Ala Gly Ala Ser Arg Val Ser Ser Pro
1 5 10 15
Glu Tyr Asn Ala Lys Lys Thr Glu Tyr Leu Gln Lys Glu Ile Arg Val
20 25 30
Lys Trp Asn Tyr Ala Pro Pro Pro Ser Asn Thr Pro Tyr Asp Pro Ala
35 40 45
Gln Ala Lys Gly Ala Ala Lys Lys Lys Lys Arg Leu Met Lys Gln Leu

```
    50                  55                  60
Lys His Arg Pro Lys Glu Gln Arg Ser Arg Gln Arg Gln Arg Gln Lys
65                  70                  75                  80
His Arg Gln Arg Ile Lys Arg Arg Pro Arg Arg Arg Gly Ser Ser Ser
                85                  90                  95
Ser Lys Ala Gly Asn Asn Asn Asn Asn Asn Arg Asn Ile Ala Phe Leu
            100                 105                 110
Asn Gly Asn Thr Gln Lys Lys Tyr Cys Trp Gln Glu Phe Arg Thr Leu
        115                 120                 125
Pro Lys Gly Met Arg Thr Asp Glu Gly Ala Gly Ala Trp Arg Asp Met
    130                 135                 140
Thr Ser Asp Gly Ala Val Thr Phe Trp Ile Phe Leu Leu Cys Leu Ile
145                 150                 155                 160
Ala Ser Pro His Leu Gln Gly Gly Val Ala Gly Arg Pro Asp Glu Leu
                165                 170                 175
His Ile Gly Gly Ile Phe Pro Ile Ala Gly Lys Gly Gly Trp Gln Gly
            180                 185                 190
Gly Gln Ala Cys Met Pro Ala Ala Arg Leu Ala Leu Asp Asp Val Asn
        195                 200                 205
Lys Gln Pro Asn Leu Leu Pro Gly Phe Lys Leu Ile Leu His Ser Asn
    210                 215                 220
Asp Ser Glu Cys Glu Pro Gly Leu Gly Ala Ser Val Met Tyr Asn Leu
225                 230                 235                 240
Leu Tyr Asn Lys Pro Gln Lys Leu Met Leu Leu Ala Gly Cys Ser Thr
                245                 250                 255
Val Cys Thr Thr Val Ala Glu Ala Ala Lys Met Trp Asn Leu Ile Val
            260                 265                 270
Leu Cys Tyr Gly Ala Ser Ser Pro Ala Leu Ser Asp Arg Lys Arg Phe
        275                 280                 285
Pro Thr Leu Phe Arg Thr His Pro Ser Ala Thr Val His Asn Pro Thr
    290                 295                 300
Arg Ile Lys Leu Met Lys Lys Phe Gly Trp Ser Arg Val Ala Ile Leu
305                 310                 315                 320
Gln Gln Ala Glu Glu Val Phe Ile Ser Thr Val Glu Asp Leu Glu Asn
                325                 330                 335
Arg Cys Met Glu Ala Gly Val Glu Ile Val Thr Arg Gln Ser Phe Leu
            340                 345                 350
Ser Asp Pro Thr Asp Ala Val Arg Asn Leu Arg Arg Gln Asp Ala Arg
        355                 360                 365
Ile Ile Val Gly Leu Phe Tyr Val Val Ala Ala Arg Arg Val Leu Cys
    370                 375                 380
Glu Met Tyr Lys Gln Gln Leu Tyr Gly Arg Ala His Val Trp Phe Phe
385                 390                 395                 400
Ile Gly Trp Tyr Glu Asp Asn Trp Tyr Glu Val Asn Leu Lys Ala Glu
                405                 410                 415
Gly Ile Thr Cys Thr Val Glu Gln Met Arg Ile Ala Ala Glu Gly His
            420                 425                 430
Leu Thr Thr Glu Ala Leu Met Trp Asn Gln Asn Asn Gln Thr Thr Ile
        435                 440                 445
Ser Gly Met Thr Ala Glu Glu Phe Arg Leu Val Leu Leu Leu Tyr Pro
    450                 455                 460
Leu Leu Lys Phe Leu Leu Gln Cys Ala Phe Phe Leu Arg Arg His Arg
465                 470                 475                 480
Leu Asn Gln Ala Leu Ile Glu Glu Gly Tyr Asp Ile Asn His Asp Arg
                485                 490                 495
Tyr Pro Glu Gly Tyr Gln Glu Ala Pro Leu Ala Tyr Asp Ala Val Trp
            500                 505                 510
Ser Val Ala Leu Ala Phe Asn Lys Thr Met Glu Arg Leu Thr Thr Gly
        515                 520                 525
Lys Lys Ser Leu Arg Asp Phe Thr Tyr Thr Asp Lys Glu Ile Ala Asp
```

```
    530                 535                 540
Glu Ile Tyr Ala Ala Met Asn Ser Thr Gln Phe Leu Gly Val Ser Gly
545                 550                 555                 560
Val Val Ala Phe Ser Ser Gln Gly Asp Arg Ile Ala Leu Thr Gln Ile
                565                 570                 575
Glu Gln Met Ile Asp Gly Lys Tyr Glu Lys Leu Gly Tyr Tyr Asp Thr
            580                 585                 590
Gln Leu Asp Asn Leu Ser Trp Leu Asn Thr Glu Gln Trp Ile Gly Gly
        595                 600                 605
Lys Lys Thr Lys His Val Thr Thr Glu Pro Asn Lys Asn Gly Asn Pro
    610                 615                 620
Ser Lys Asp Asp Val Asn Thr Arg Gly Ser Leu Thr Ser Leu Ser Glu
625                 630                 635                 640
Lys Asn Asp Glu Ser Lys Asn Pro Ser Ser Ser Ala Ser Leu Gln Gln
                645                 650                 655
Ala Pro Lys Val Ala Pro Lys Pro Lys Lys Leu Ser Ile Ser Asp Ala
            660                 665                 670
Gly Lys Asp Thr Val Thr Gln Lys Val Lys Glu Asn Glu Glu Pro Gly
        675                 680                 685
Phe Ser Lys Gln Phe Glu Lys Glu Ser Ile Gly Val Arg Ala Asn Arg
    690                 695                 700
Asn Ser Ser Ala Thr Lys Lys Glu Asn Glu Lys Leu Leu Val Lys Thr
705                 710                 715                 720
Val Pro Gly Lys Ser Leu Ile Lys Glu Ser Asn Asp Glu Asn Val Glu
                725                 730                 735
Pro Ser Arg Arg Thr Lys Ser Gln Pro Val Gly Lys Lys Val Tyr Glu
            740                 745                 750
Glu Ser Ser Thr Arg Arg Val Arg Glu Pro Phe Asp Ser Phe Asp Arg
        755                 760                 765
Glu Lys Tyr Leu Ser Asp Met Ile Asp Tyr Asp Arg Ser Ser Glu Asp
    770                 775                 780
Val Pro Glu Lys Lys Leu Ser Trp Glu Pro Asp Ser Thr Leu Arg Arg
785                 790                 795                 800
Arg Phe Val Ser Asn Asn Glu Tyr His Ser Ser Leu Glu Gly Glu Glu
                805                 810                 815
Glu Glu Ile Asp Leu Asp Ser Val Gly Ser Ser Thr Met Arg Gly Ser
            820                 825                 830
Tyr Arg Met Pro Gln Asn Ala Glu Asp Glu Arg Pro Val Ile Leu Ala
        835                 840                 845
Glu Ile Ile Asn Met Gly Lys Leu Lys Glu Leu Lys Glu Gln Arg Glu
    850                 855                 860
Ser Thr Asn Val Phe Arg Asn Asn Arg Arg Asp Asp Lys Thr Ile Asp
865                 870                 875                 880
Glu Val Lys Gly Ser Lys Ile Thr Gly Thr Ser Lys Gln His Val Glu
                885                 890                 895
Glu Val Lys Tyr Asn Glu Thr Ala Ser Asn Glu Asp Asp Asp Asn Ala
            900                 905                 910
Gln Val Lys His Ile Gly Thr Ile Thr Lys Ile Phe Ser Ile Ala Gln
        915                 920                 925
Arg Met Lys Glu Lys Lys Ser Ser Lys Glu Ser Glu Glu Asp Glu Asn
    930                 935                 940
Lys Asp Glu Asn Lys Asp Val Lys Lys Asp Glu Asn Lys Asp Asp Glu
945                 950                 955                 960
Gln Lys Leu Pro Pro Glu Pro Glu Ile Glu Leu Glu Ala Lys Lys Ala
                965                 970                 975
Trp Thr Phe Pro Ile Ala Glu Thr Ser Pro Asp Ile Pro Gly Glu Asn
            980                 985                 990
Gln Ile Leu Gln Glu Tyr Gln Lys Val Lys Ile Gly Lys Arg Asn Arg
        995                 1000                1005
Lys Tyr Arg Cys Ile Gly Ile Asn Thr Asp Ile Ser Arg Lys Ser Lys
```

```
    1010                1015                1020
Ile Lys Thr Phe Ile Ser Lys Ser Ala Gly Asp Asp Leu Val Val Tyr
1025                1030                1035                1040
His Asp Gly Gly Arg Leu Arg Asp Ile Gly Gln Thr Thr Glu Asn Leu
                1045                1050                1055
Lys Asn Gln Arg Lys Asn His Lys His Lys Thr Glu Pro Asp Asp Ile
            1060                1065                1070
Pro Val Asp Ile Gly His Ser Asp Asp Arg Val Arg Glu Ile Gly Val
        1075                1080                1085
Asn Thr Lys Lys Leu Pro Lys Ile Ile Ile Pro Pro Ile Ala Glu Met
    1090                1095                1100
His Val His Lys Asn Gly Lys Leu Arg Asp Ile Gly Thr Ser Thr Asp
1105                1110                1115                1120
Lys Pro Phe Trp Pro Ile Asp Asp Gly Thr Asp Val Ile Tyr Met His
                1125                1130                1135
Pro Ile Lys Thr Asp Arg Lys Lys Leu Asn Lys Leu Ile Val Asp Pro
            1140                1145                1150
Pro Pro Asp Asn Gly Pro Tyr Lys Met Pro Thr Lys Glu Asp Arg Arg
        1155                1160                1165
Thr Tyr Tyr Lys Gly Cys Glu Tyr His Phe Pro Gly Arg Thr Glu Trp
    1170                1175                1180
Arg Arg Leu Phe Phe Asn Lys Ile His Gly Lys Tyr Lys Leu Arg Arg
1185                1190                1195                1200
Pro Ser His Trp Leu Tyr Thr Leu Val Phe Ser Val Leu Tyr Ile Leu
                1205                1210                1215
Phe Val Ile Ile Phe Ser Met Ala Trp Phe Asp Phe Ile Lys Asp Asp
            1220                1225                1230
Ala Ser Arg Lys Val Pro Met Ile Lys Met Ala Gln Pro Phe Ile Ser
        1235                1240                1245
Phe Thr Pro Ile Gly Pro Arg Thr Asn Pro Lys Ala Val Ser Phe Asp
    1250                1255                1260
Pro Arg Asn Ser Thr Glu Val Met Glu Lys Tyr Ala Gly Ile Met Ala
1265                1270                1275                1280
Leu Leu Glu Lys Tyr Gly Asp Tyr Gly His Asn Pro Arg Phe Gly Thr
                1285                1290                1295
Cys Thr Ala Asn Glu Lys Phe Gly Tyr Pro Ser Gly Glu Pro Cys Val
            1300                1305                1310
Phe Leu Lys Val Asn Arg Ile Ile Gly Phe Lys Thr Glu Pro Tyr Ile
        1315                1320                1325
Asn Ser Asp Glu Leu Val Lys Ala Lys Ile Asp Glu Val Glu Phe Thr
    1330                1335                1340
Ala Leu Lys Arg Leu Leu Glu Asn Thr Thr Thr Glu Glu Gly His Leu
1345                1350                1355                1360
Asn Arg Thr Trp Ile Thr Cys Arg Ser Asp Lys Asp Lys Asn Val Leu
                1365                1370                1375
Ile Glu Phe His Pro Glu Pro Ala Ile Arg Thr Glu Tyr Thr Asp Ile
            1380                1385                1390
Glu Glu Lys Ile Glu Tyr Ile Ala Asn Glu Gly Lys Lys Ser Phe Phe
        1395                1400                1405
Gly Pro Asn Asp Val Asn Arg Ile Val Ala Leu Lys Ile Lys Asn Leu
    1410                1415                1420
Lys Ala Asn Glu Arg Val His Ile Asn Cys Lys Ile Val Ile Gln Ser
1425                1430                1435                1440
Ser His Pro Val Cys Asn Thr Ile Met Leu Phe Gly Val Ile Ile Cys
                1445                1450                1455
Leu Ile Ser Val Ile Leu Leu Gly Ile Asp Gly Arg Phe Val Ser Pro
            1460                1465                1470
Glu Glu Tyr Pro Lys Ile Phe Asp Pro Leu Gln Arg Tyr Leu Glu Thr
        1475                1480                1485
Phe Pro Leu Glu Asp Pro Val Ser Thr Thr Asp Asp Ile Lys Ile Arg
```

```
    1490                1495                1500
Pro Glu Leu Glu His Cys Glu Ser Gln Arg Asn Ser Met Trp Leu Gly
1505                1510                1515                1520
Leu Val Tyr Gly Phe Lys Gly Leu Ile Leu Val Phe Gly Leu Phe Leu
                1525                1530                1535
Ala Tyr Glu Thr Arg Ser Ile Lys Val Lys Gln Ile Asn Asp Ser Arg
            1540                1545                1550
Tyr Val Gly Met Ser Ile Tyr Asn Val Val Val Leu Cys Leu Ile Thr
        1555                1560                1565
Ala Pro Val Gly Met Val Ile Ala Ser Gln Gln Asp Ala Ser Phe Ala
    1570                1575                1580
Phe Val Ala Leu Ala Val Ile Phe Cys Cys Phe Leu Ser Met Leu Leu
1585                1590                1595                1600
Ile Phe Val Pro Lys Val Ile Glu Val Ile Arg His Pro Lys Asp Lys
                1605                1610                1615
Ala Glu Ser Lys Tyr Asn Pro Asp Ser Ala Ile Ser Lys Glu Asp Glu
            1620                1625                1630
Glu Arg Tyr Gln Lys Leu Val Thr Glu Asn Glu Glu Leu Gln Arg Leu
        1635                1640                1645
Ile Thr Gln Lys Glu Glu Lys Ile Arg Val Leu Arg Gln Arg Leu Val
    1650                1655                1660
Glu Arg Gly Asp Ala Lys Gly Thr Glu Leu Asn Gly Ala Thr Gly Val
1665                1670                1675                1680
Ala Ser Ala Ala Val Ala Thr Thr Ser Gln Pro Ala Ser Leu Ile Asn
                1685                1690                1695
Ser Ser Ala His Ala Thr Pro Ala Ala Thr Leu Ala Ile Thr Gln Gly
            1700                1705                1710
Glu


<210> 178
<211> 4463
<212> DNA
<213> Drosophila

<400> 178
tgtgggtggt tcgcacgccc gccatttccg gcgcggcctc ggataacgtc ctcgattacg 60
ccctgccgcc caacacgccc actccgaagg actacgcgca caagttcctg cccgagggca 120
aggatgtggt acccggaagc tcttacatac cgctgcggtt gctggtgata aagcgctgaa 180
gctgatatat gaatcttggg cggcggcttg ccttttgttg ttcgaaaaca agttccggcg 240
gaatgccgat acccgacact taatgcccaa ccagtagttg gctaattggc tttaagctta 300
ctggttccta gcgctagtct agcaatagtt tagttgttta tgttaaggcc caagttcgct 360
aatctaaatt aagccaccac atcccggggt ccgtgactaa gacttccgac cttatcgcat 420
ccaataaatc caaaactaag actcagccac atgccgtcgt aatctatgaa atacattcac 480
caaaaactaa gttgtcatct catttacata aattagaagt cataaatttg aagtgagcgg 540
ggaactactt ggctcgaaac ggttagaagc tgaattcaaa accggttaac gtttatctgt 600
cagctgtaat caaaagctta aagcttcgct tgtttggttt agtctttgac taagtttgga 660
aagcaggact aaaagctttt atttgccgaa actaaagctc ggaggtattt atttattcac 720
aagtaaaatg acaaaacttg gaagttcttt ggtaatgtaa tgaacgccaa tatttacatt 780
cgggttttat tgtatttgaa aagtctgatg aatcattatc ttgatgactg gcccctaatt 840
tacttttcat gttctacgca aggggtatga taactggttg acatggatac caagattaga 900
gcttgcagac tttataatca tccaacaaac attatggtgc attgtatacc atttatgtta 960
tgaatgaaag tggttcagat aagttacatt ttattttatc tcacacatct aagtttttgt 1020
attgctgcac aggaaaataa atcaagatta tccacaatag tttaaaagtc aatttatagg 1080
tactatgtag aatttattta cattttgttg tactaatggt aacttaactt aggtttttaa 1140
atactagtac tgtttgttga tgaatggctt aggccttgca aaatcgagtc agtgggcatt 1200
tgtatagttt gctgatttct cttcttgctc gaattgcaac agagacccaa ccaggcccgc 1260
tggttctctt tgttaaatac tataaagtaa aggaataaaa caaatccctg taaagtggct 1320
gtcaggcaga ataaataggc gaagatcctt ccgaaatcga agtatgtaca caatccaaaa 1380
```

atccatgtta ttcccaaaag aaagaagagc aacacaaaca ggccgagttg cttgaagata 1440
aggtcacgtt tgaacagggc tgtatagaca ctccaggaga tgtatcccac caatattgca 1500
ttggccaccg tgataaggcc aatagggagt atgacgccca gacttaaacc atatccagaa 1560
ggatagcaga ggatgggata gtccattgag gaattattgg gtctgtagga accgggctcc 1620
agaaaaacca ccaacaaggt gggtataagt ggcaatgtcc aggccgccac agccgacatt 1680
agtatataat gtcgtgggtg attaactcca atgactctca catatctttg atactgcaga 1740
aatccaatga taaacatcca gctgaagaca accaagagta gatattgcat cacggctcca 1800
accagagtac atcgttccga ctcagattgt tccagctgct ccagcaaatg gctttgactc 1860
aggatcagga agaagagcaa ctgaagaccc agggcagcac acagattcag caatatcttc 1920
gtggaggcta atgtccggaa gctcttaaac acggctgcag ttataaaaat cattaggagg 1980
cccaacagag aaagggtcaa gccgacattg gtgattacat ctaaggagtg atcatcttca 2040
ttggtgtcca gacctgcctg cattttgctc actcccagta gaaatgtaaa ctgggtcaga 2100
tgatccgcat ggcagaggat cactgggtgc gatgaggctt ccaagctgcc accaccgctg 2160
gtggatattc catcgctaag ccaggtctca tagttccagt aaccacatcc tccctcaaat 2220
tgagtttcat tcgcttttgc attacgcaag aaaaatggca actttccagg tagcttggta 2280
tcttacgaaa agattttaga ttaaaacata tatgctataa ttttccctca acccactcac 2340
cattcaatcc tggtattgtt atggatataa cattactagt tggccgtcgt ttgatttcct 2400
cttccgtctc cacaaagagt gaatctcgag tatagacctt aaaaactaag taagaggctc 2460
cctttctctt cagttgggtc catagttttt cgggaataaa gaccgcactt tctagattct 2520
tcatggccct tatatcttcc tcattatccg ataggtgaag catttcgaaa tgcgacatat 2580
tgccaccaga gaagttcact atcccactca catttctcac ttctggatct acaaagaata 2640
ctctgagttc gctcgtcttt agaacaatga cacccatatc aaacaaggga tggaatgtga 2700
actcagtgct gtttgcaaaa atagtggtgt ttaactttaa aattgttgga ctaagctcat 2760
ccaaataatc ctcgaagttg ctaaggagtg cattggttgc attagcgtct gctgaaacac 2820
gaagggtttc attatctgta gacataactt tttggcatat cttcatcaag tcccctacaa 2880
ctgcagcatt tctttccggt tttttgacca gcgaatcgaa tatctccgaa atactgaaga 2940
catcggcggc ggccagtttc tcctgagttt gagtcaggag gttactcagt tggctaatca 3000
tgtccggagg gctcttcaca gcgagaaggt ttatactcga gctgagttcc ggtgccgacc 3060
tacaggtcac gtgatccaag ctctgccaaa ctctattccc attgcgaatt tggcacctcc 3120
gtcggatggg aattccccgt tcgttcaggc atagatccct tagattcgcc ttttgcccaa 3180
ctctcgtttt ctccacgcga ttgattcgct tttccttgcc ttcatccgtt tcgtattcca 3240
gttcaaagtc ctcggccttg cagaatcctc ggttgcagtc cttgtaggga ttgtgcgaca 3300
gctgaccatc tgcatccacg cagaggagga agcgcaagtt cttcaggtgg cagtagtcca 3360
gacaccgggc cacattcggc tgatccaggg cgtacagggg gggatttgta cctccagata 3420
cgagtattat atagccagca atccaccaga gaaacagcag catactgcta tctattttta 3480
actaatttaa gcgcaggctg cgatgtcgca ctcccgaagc tccgcgatcg actgacgagg 3540
cggagataaa atgcgagtgg cgataaaacg agttttcctc tcctcatccc tgtaaatagg 3600
gcctgcaaga attgccactt tcctcactgt tctagctggg tatttgaatt gctggtctac 3660
ttcgagcaga gctcaaggtc aactggttat ctggccgacg ataagaattg cagcatcttc 3720
aaaggtcgaa cggtgcgttc ttttcaaaat atctttgaat ataataagtg aaagtctcga 3780
aaataacacg aacctcttta ataatgggta ttatttatac acagtagttc tgtttacttt 3840
gatcacattt tttatcaagg gaacttccta attaaaaaca aaaaaaaaaa aacgaatagc 3900
ctgataaaga gctaaatggt aattatcaaa attatatata ccatacaccc aacgataatg 3960
ataatttgtt tatcgcagta aagtgagatt ctctgaattg ataaaggtta attgccgttt 4020
ttcaacacca ttgcatgtat cgataaaact aataggaatt acaaataaat taaaaatctg 4080
aattttcttt tttgctctgg gtaacaatgg ctgccatttc tgaatgtttc tgttattctc 4140
tctgatttta aaagtgcaac gatacttaac acgaactttg agttgctagt cgccataagc 4200
atatgaatgc atcaatcggt agctaggggc tgtcaacaac tacaatatgt atgataagca 4260
gccaccttta gatactttcc ttactaacac ctagaaggta gagatagagg tattcaaata 4320
tttgcattca gatataaaat aatgttctct tgttcaagct gtgaaatctt tatctgtaga 4380
gataaagtga gcatttcaat tgaagaacca atttgtatat ttttttagat ttcaccgaac 4440
gcatgaatag aatttgtctg cta 4463

<210> 179
<211> 1938
<212> DNA
<213> Drosophila

<400> 179
atgctgctgt ttctctggtg gattgctggc tatataatac tcgtatctgg aggtacaaat 60

```
cccccccctgt acgccctgga tcagccgaat gtggcccggt gtctggacta ctgccacctg 120
aagaacttgc gcttcctcct ctgcgtggat gcagatggtc agctgtcgca caatccctac 180
aaggactgca accgaggatt ctgcaaggcc gaggactttg aactggaata cgaaacggat 240
gaaggcaagg aaaagcgaat caatcgcgtg gagaaaacga gagttgggca aaaggcgaat 300
ctaagggatc tatgcctgaa cgaacgggga attcccatcc gacggaggtg ccaaattcgc 360
aatgggaata gagtttggca gagcttggat cacgtgacct gtaggtcggc accggaactc 420
agctcgagta taaaccttct cgctgtgaag agccctccgg acatgattag ccaactgagt 480
aacctcctga ctcaaactca ggagaaactg gccgccgccg atgtcttcag tatttcggag 540
atattcgatt cgctggtcaa aaaaccggaa agaaatgctg cagttgtagg ggacttgatg 600
aagatatgcc aaaaagttat gtctacagat aatgaaaccc ttcgtgtttc agcagacgct 660
aatgcaacca atgcactcct tagcaacttc gaggattatt tggatgagct tagtccaaca 720
atttaaaagt taaacaccac tatttttgca aacagcactg agttcacatt ccatcccttg 780
tttgatatgg gtgtcattgt tctaaagacg agcgaactca gagtattctt tgtagatcca 840
gaagtgagaa atgtgagtgg gatagtgaac ttctctggtg gcaatatgtc gcatttcgaa 900
atgcttcacc tatcggataa tgaggaagat ataagggcca tgaagaatct agaaagtgcg 960
gtctttattc ccgaaaaact atggacccaa ctgaagagaa agggagcctc ttacttagtt 1020
tttaaggtct atactcgaga ttcactcttt gtggagacgg aagaggaaat caaacgacgg 1080
ccaactagta atgttatatc cataacaata ccaggattga atgataccaa gctacctgga 1140
aagttgccat ttttcttgcg taatgcaaaa gcgaatgaaa ctcaatttga gggaggatgt 1200
ggttactgga actatgagac ctggcttagc gatggaatat ccaccagcgg tggtggcagc 1260
ttggaagcct catcgcaccc agtgatcctc tgccatgcgg atcatctgac ccagtttaca 1320
tttctactgg gagtgagcaa aatgcaggca ggtctggaca ccaatgaaga tgatcactcc 1380
ttagatgtaa tcaccaatgt cggcttgacc ctttctctgt tgggcctcct aatgattttt 1440
ataactgcag ccgtgtttaa gagcttccgg acattagcct ccacgaagat attgctgaat 1500
ctgtgtgctg ccctgggtct tcagttgctc ttcttcctga tcctgagtca aagccatttg 1560
ctggagcagc tggaacaatc tgagtcggaa cgatgtactc tggttggagc cgtgatgcaa 1620
tatctactct tggttgtctt cagctggatg tttatcattg gatttctgca gtatcaaaga 1680
tatgtgagag tcattggagt taatcaccca cgacattata tactaatgtc ggctgtggcg 1740
gcctggacat tgccacttat acccaccttg ttggtggttt ttctggagcc cggttcctac 1800
agacccaata attcctcaat ggactatccc atcctctgct atccttctgg atatggttta 1860
agtctgggcg tcatactccc tattggcctt atcacggtgg ccaatgcaat attgcaatac 1920
aaaaacttag atgtgtga 1938
```

<210> 180
<211> 645
<212> PRT
<213> Drosophila

<400> 180

```
Met Leu Leu Phe Leu Trp Trp Ile Ala Gly Tyr Ile Ile Leu Val Ser
 1               5                  10                  15
Gly Gly Thr Asn Pro Pro Leu Tyr Ala Leu Asp Gln Pro Asn Val Ala
            20                  25                  30
Arg Cys Leu Asp Tyr Cys His Leu Lys Asn Leu Arg Phe Leu Leu Cys
        35                  40                  45
Val Asp Ala Asp Gly Gln Leu Ser His Asn Pro Tyr Lys Asp Cys Asn
    50                  55                  60
Arg Gly Phe Cys Lys Ala Glu Asp Phe Glu Leu Glu Tyr Glu Thr Asp
65                  70                  75                  80
Glu Gly Lys Glu Lys Arg Ile Asn Arg Val Glu Lys Thr Arg Val Gly
                85                  90                  95
Gln Lys Ala Asn Leu Arg Asp Leu Cys Leu Asn Glu Arg Gly Ile Pro
            100                 105                 110
Ile Arg Arg Arg Cys Gln Ile Arg Asn Gly Asn Arg Val Trp Gln Ser
        115                 120                 125
Leu Asp His Val Thr Cys Arg Ser Ala Pro Glu Leu Ser Ser Ser Ile
    130                 135                 140
Asn Leu Leu Ala Val Lys Ser Pro Pro Asp Met Ile Ser Gln Leu Ser
145                 150                 155                 160
Asn Leu Leu Thr Gln Thr Gln Glu Lys Leu Ala Ala Ala Asp Val Phe
```

```
                165                 170                 175
Ser Ile Ser Glu Ile Phe Asp Ser Leu Val Lys Lys Pro Glu Arg Asn
            180                 185                 190
Ala Ala Val Val Gly Asp Leu Met Lys Ile Cys Gln Lys Val Met Ser
        195                 200                 205
Thr Asp Asn Glu Thr Leu Arg Val Ser Ala Asp Ala Asn Ala Thr Asn
    210                 215                 220
Ala Leu Leu Ser Asn Phe Glu Asp Tyr Leu Asp Glu Leu Ser Pro Thr
225                 230                 235                 240
Ile Leu Lys Leu Asn Thr Thr Ile Phe Ala Asn Ser Thr Glu Phe Thr
                245                 250                 255
Phe His Pro Leu Phe Asp Met Gly Val Ile Val Leu Lys Thr Ser Glu
            260                 265                 270
Leu Arg Val Phe Phe Val Asp Pro Glu Val Arg Asn Val Ser Gly Ile
        275                 280                 285
Val Asn Phe Ser Gly Gly Asn Met Ser His Phe Glu Met Leu His Leu
    290                 295                 300
Ser Asp Asn Glu Glu Asp Ile Arg Ala Met Lys Asn Leu Glu Ser Ala
305                 310                 315                 320
Val Phe Ile Pro Glu Lys Leu Trp Thr Gln Leu Lys Arg Lys Gly Ala
                325                 330                 335
Ser Tyr Leu Val Phe Lys Val Tyr Thr Arg Asp Ser Leu Phe Val Glu
            340                 345                 350
Thr Glu Glu Glu Ile Lys Arg Arg Pro Thr Ser Asn Val Ile Ser Ile
        355                 360                 365
Thr Ile Pro Gly Leu Asn Asp Thr Lys Leu Pro Gly Lys Leu Pro Phe
    370                 375                 380
Phe Leu Arg Asn Ala Lys Ala Asn Glu Thr Gln Phe Glu Gly Gly Cys
385                 390                 395                 400
Gly Tyr Trp Asn Tyr Glu Thr Trp Leu Ser Asp Gly Ile Ser Thr Ser
                405                 410                 415
Gly Gly Gly Ser Leu Glu Ala Ser Ser His Pro Val Ile Leu Cys His
            420                 425                 430
Ala Asp His Leu Thr Gln Phe Thr Phe Leu Leu Gly Val Ser Lys Met
        435                 440                 445
Gln Ala Gly Leu Asp Thr Asn Glu Asp Asp His Ser Leu Asp Val Ile
    450                 455                 460
Thr Asn Val Gly Leu Thr Leu Ser Leu Leu Gly Leu Leu Met Ile Phe
465                 470                 475                 480
Ile Thr Ala Ala Val Phe Lys Ser Phe Arg Thr Leu Ala Ser Thr Lys
                485                 490                 495
Ile Leu Leu Asn Leu Cys Ala Ala Leu Gly Leu Gln Leu Leu Phe Phe
            500                 505                 510
Leu Ile Leu Ser Gln Ser His Leu Leu Glu Gln Leu Glu Gln Ser Glu
        515                 520                 525
Ser Glu Arg Cys Thr Leu Val Gly Ala Val Met Gln Tyr Leu Leu Leu
    530                 535                 540
Val Val Phe Ser Trp Met Phe Ile Ile Gly Phe Leu Gln Tyr Gln Arg
545                 550                 555                 560
Tyr Val Arg Val Ile Gly Val Asn His Pro Arg His Tyr Ile Leu Met
                565                 570                 575
Ser Ala Val Ala Ala Trp Thr Leu Pro Leu Ile Pro Thr Leu Leu Val
            580                 585                 590
Val Phe Leu Glu Pro Gly Ser Tyr Arg Pro Asn Asn Ser Ser Met Asp
        595                 600                 605
Tyr Pro Ile Leu Cys Tyr Pro Ser Gly Tyr Gly Leu Ser Leu Gly Val
    610                 615                 620
Ile Leu Pro Ile Gly Leu Ile Thr Val Ala Asn Ala Ile Leu Gln Tyr
625                 630                 635                 640
Lys Asn Leu Asp Val
```

645

<210> 181
<211> 6938
<212> DNA
<213> Drosophila

<400> 181
cagctagtag tgaataatac ttgagttttt gagtttgttg ggctgggttg aaaaaaaaaa 60
atatgcactt tagcctctat taagcatttt cctatggttt tcactctcgt tcgggaactg 120
tgtgttttga aaatgtaaca atttgcgctt tcgtcctcgt acattgtgtg tatttacaga 180
aagtctagcc cattttttaga ttttgtatgg gaatgagtgg aaattacaag tgtataatat 240
ataatacaaa tacactttag ttggtaagat tagcaagacg taagacgtac gatagttaat 300
gaggggggaa ttgctagttt tatgttaatg cctataagct caaatttaca actgattatt 360
cgctgatcga aaatgatctc aagatctcta tcatttgttt agctttagtg tggatagtgt 420
tgaaaggttg acttgaaagc aaatggttag taactaactg tcgaggaact tagaggtaaa 480
cgaaaacgcg tgttttagga atgggtttgt ttttcttagt gtatgtggcc ttgtttttac 540
agttatatcg ttacctacac aggtgggctc tgttttaccg atctaaacgt ttgaacaact 600
ttgatctatc agtgatccta atcaatcgaa atcaatcgaa ctcaagactg acagaaatca 660
acgtaaacta aagtaaaatc gtctaaaaat ttatcaaaac tttataaaac ttcttggctg 720
ctttcaaggt tcctcacttt acgtttcatt gaaaaataaa gtaatatttg ttattgtttg 780
tgtttggtga gtgtttgttg catctccatt tgcatcaatc attttcctgc tggttttttt 840
tttctttctt tgcgaaatta aagtaattct tttattgttt gtgttgttgc gctatataat 900
aactatttct cacatattat ctacacatac caacgtgtat cgtaccgttt aatacctata 960
tacaattctc aaatatccgc catgaaccag caaggactcc ctacttcgaa tgcgagacca 1020
cggagtggga gaccatcgag tgagagagca ttgagtggga cacgccggag atttccgcat 1080
aggtatccgt gggtccttcg gcactgcaca ccgaatcctc gggcgtcctg gaaggttccg 1140
cctcggtgat ctcgatgtcc tgatgatgct tgttgctgcg gtgtttggca cttagagcct 1200
gcgaatttga gatattttag tgggtcacat ggacaactgg tggatatgtg ttttctcgtt 1260
gtagtttcgt gtgcaatctg ttttcgataa gacttatgtg gtgctatcta tgggacggac 1320
acatgcacac acacacacac acacgcacaa gctggtggtt tcgggttcat tggggaactt 1380
tgggtctaag tttatagaca aacgtgcgag aaattctttt gatttagtcg ttttgttttg 1440
gattttagtt tacattcaca acacagatat tttagcaggt taggtttaga ggtttatgct 1500
aaaagcaacc gaaaattctt accttatcct tgctgccctt tttctacatt gaacattggt 1560
ggttattgtt ttgtttggtt tggttgtggt gggtttcgaa tatattgtac atagtcatga 1620
caagaacaat gatatttata gatgtattaa tataagtggc aaacagtact gcgagttcac 1680
attggaggat caagagaagt acagattcat agacggttag agagacgggg gagagagaaa 1740
gatctaggga gcgtggggat gggcagcaaa aggtttggtt tcggattgaa gaataacagc 1800
aatatcgatc tgaggcctac agcccggttc cccattctgg ccacgtcaat gtcatgtaca 1860
gtgagcacta tggctggtgg tggactacgc tgccgttgtg gtgaccccga ctcacctgca 1920
gggagaagcg acggtgaccc gccttgcgtc cgccacgtcg cttactgatg tgcggattgt 1980
cctgccgcag agtcttgttc ttcattatct ccaattgggt atacagtctg gaatcgaaaa 2040
tgtatagaaa atagagataa tagaacatat agatatgctt cgtacctttt gagttcggcg 2100
cgtatgtcct cgggactcat ttcggatatg gtcagctcac caatgtccgg atcgcccaga 2160
cagagcccgg catatcctcc gcccagccgc tggccggcgt cgtgtgacgt acccttgaaa 2220
gcatccacag gtaaacgtat tgatagatca tacgctagcg aacgaacctg caagctatga 2280
agatgcaaca aacaccgcaa accacactca ggaagtcaaa cggaaaatgg agaacggaac 2340
ggaaacggaa atggaaacgg aactcaaagg taagagagag caaaatcgca tggtatagat 2400
ctgtgctgaa gagttatgta tagttatatg tatatagtat gtgtatgggt ggtctgtgct 2460
tattggtcgt gattgtatct ttgggggtgc tggaaatgaa cgaaaagaaa caaatcaaat 2520
taaaaaagta actcaaaagg taaacaagat ggacgatagc tataagaaga ctacttagac 2580
agagaacttc agtttaagat tattcagttg aggcacagaa agacagagaa agttacagtg 2640
aggcagagtc ttggatacac tagtttggaa ttttttttc tggtatgcat attggaaatt 2700
ttgaaaggaa accgattttt gtcatttgtt tgttagtagt tagatgttgt gttttcttgt 2760
taaattcatg caaaaatcaa cactaaacat gcaaattaaa tggcgtgctt attgagttca 2820
agttaaagtg caatacattc aggcataaat aataagagct aaaaacagtc gaacattctc 2880
tgtttaaacg agacacactt gcagcaactt catagcttga ctaaagtttc aattcaatgt 2940
tagcttatta ctaaaattat gttaattttc tcggtgaaaa ccagaacaag aactcaaaat 3000
ataaccataa acaaatgaaa tattcaaatt gtattttaat tagcaaaagc aaccaagctt 3060

```
ttgagtggat gtccacattt taatacttct tgtgttttgg ctagttattt aacaaaaatc 3120
gtttcctcaa ccaatgtcaa taccaagtac gcatttcgat gactataatg acgaccacgc 3180
ccacgacaga cgaaaggcac aaaatgaacg ggtgtgaata ctctaagcat atgaattata 3240
ttcgtttggt gcccgactag gacaaaacac ttcctaatta catcatatat ctcaatgtgt 3300
agaacacaaa agcgtactca aaccaaaaaa aaaaaaaaaa aaaccaaagc aaagtaagta 3360
aatcatatcg tttcgaactc aaatcggtta cctgcttgtg ctggtaccac aactttggca 3420
caaatatcag acccaaggcg aaactattcg tcagctggga gcggatgaac agggccagca 3480
agatggcgct gggactcatt tccggcaggt agacaaagcg cagaaagtag aagctggacg 3540
agaccaggaa ctccagcgtc agtgcggtga ccagaaattg tcgttccttc aatgggacaa 3600
aaggaaatta ccgattaatt catttgactt gacttaaatt ccccttctac ttagttaatc 3660
atctaagcca atctcttgcg cttaattgat aaagaaagat ctgcaatcga tttgactgta 3720
catcttaccc ggaactgggt gttggcattc cgactggcga tggacaggtg cagtccaaag 3780
cacaggatga gcatctcgct ggtctgcgtg accagttccc acttgagcgg atggcaggtg 3840
ttcgtgtccg cctccctgag gctctccagc tgcgcacttt cgagaagatc caacgacgag 3900
gccgtaaagg cggccatgta gcagatgacg gcaaagacca tggtgcccag atacttgagc 3960
aggtccacgt cgcgcagcac ccagcgatgc gctttccggg tgcgaaagtc caccaggtgg 4020
cggtacaact tcagtatgat ggcgccgtag caggtgatga agcccagctc ccgaagccag 4080
ggctccagaa ggcagcgctc ggtggatgcg ggaaagaaat ggacggcaac ctaaagatga 4140
atatgtaaat acacacagaa gttgaaggtg accatcgcca ggagcactca cagatgcata 4200
aagtaaaaca atgcccagca gtatcgtctc cagcacagtc cacataccag acgcaatggc 4260
ctgaaaatga tgaaaggatt actcaaagtt tcccaacttt aggatatgta aactaaatac 4320
cttgcacttt cgctgccgga agacaatcac cccgaggaca atgcagcaca ggatgcaggc 4380
gcccagtacg atggccacca ggaggcgcag acaggcgtcc acattgagca cctcctcctc 4440
ctgaaaggtc agacagacgc cgttgctatc gcagttagtg catccgccgg gacacggaat 4500
gcacgagtag ttgtcgtagc cctcggacag ttccactcga tcgccacgga atccctggag 4560
cgtggaattg ggcaggtagt aggattcccg gcacagacag gtgtacacat cccgggtggc 4620
ggcgggtttg ttctcggtga gcaggcagaa ggtcgtattc cgatcgcaac tacagagata 4680
tggtattata tatattatcc agaactaatg tcacgaatta tcgtataaag tttggccaca 4740
actcacccat gacggcgacc gaaaacttcc tccaggccat cgttgcacac atcctcgtcg 4800
gcggcaatga aggcagccgc cacaactctg aaagcattta gcataaatta caaagttgtg 4860
ccaaccagct gcacatcgat gcgactcact tgatcctgtg ctcgctgaag gatatgcgaa 4920
atggccagag ccagcgacgc gtcaacaggt tgcagtcgcg aaatggagcg ccccactcct 4980
gggtgccgtc cgtgctgttc caggccgcac tgtagtcggg taaaataata catttatggg 5040
caactttcca agattatacg gcttatggtc tattcaaatt aagcgatgca cattatttcg 5100
gcaagctacg gtgcatattg aatttaattg tatttatctt ttagagtaat tagtatttgc 5160
atttcaaagc cataagataa ttgattttaa tttgttgcta attaatgtcg cacacttttt 5220
tgggacacaa ccagtcttcg aaacttaccc caactcccgc caataggaac ccaaatatgc 5280
cttattgtgc acctgatccg gtggtatctg caataagtta aatcgattag cccatggccc 5340
cattatcgct gattagcgag tgcataccgt aacgttatcc ctgaatttaa cgaccaccgt 5400
gtgaatgctg ggcgcggcaa tggcgaatgc cagcagattg tcgtcctcgt tgagggcgta 5460
gctggcgatt ccattggccc ggcgactcac ctccgcacgt gtcggctcat cgatggcaat 5520
cggtgtgttg gcatttggga atttgcccga taaaaattca ataaactggc gtgcggcatc 5580
ctcggcgatg tcggtgtagt cgtaaaactg ttgaattcat aaaaccagat cacaattaat 5640
tgcgttgtgt cataaataat catttcaatg tatatgaaat gggcatataa atatcattta 5700
aatcgaattc gatttaaata cacaaagaaa aagttcaata aaattaacac agcataaatt 5760
tacgctttgc taaactcaat ttatttatac cacactatat aggaaacgga aaacgtgtgc 5820
tttctaattg aagatatttt cggaaattgc gaaaaaatta cgtatatgcc acgtatattg 5880
gtcgccacgt gatttaagtg accatgcggc acggccactt cccccgactt gggaccatta 5940
attattcacc tggcgcacag gacacgtttc acagcaacac cttgcggcaa gagtgcgaaa 6000
agttttgctt cgtttcgact gtgaaataat aggatttaat accctcgccg tgctgtggga 6060
ctgcgggaag ggtttttaat ccaagttcgt tcagctccta accgaccaaa tgtttataca 6120
taatttacac gttacaagaa actctcgaga tctacgcacc aaagtgtagg aaatggtagg 6180
gccactttga tgacatcatt atggctttgg gaaacaaaag taaccacaga agggagcagg 6240
gagggggttg ctggtagcgg ggaaaaagga cacgttgttt gggcgtgagt catgtagttg 6300
ggctggcggc acctttggc ttttaaaaca aatatttccc attagctggg ctttcggtga 6360
tgagcgacca cattgcgtat acgccaccga agccaccaaa taaattgcca acaagtaaca 6420
tcgatataca aacgtttgct ggatgctcct ctaccatgtg tgtgtgagtg tgtttcttca 6480
ttttttttcc tgccaaacgt gagccaaaat ttacacatcg aggcaaaagc aatcagtata 6540
aggcagctgg agaaatggca cgagacctgc tcccagataa agtcattatg tagtgctcct 6600
agcactcgcg cgcagccata atcataaaca tttggtatac tcttgcatgc tagagaaata 6660
```

ttgtagttgg gtaagtgttt tcagatatta aactaagtaa ggaaatattt ctatttttac 6720
accattattc attgtttttt cttcaggagt tatctgaact ttaagtaatc ccaccacagt 6780
tcttggaagc agtagaatgt cttcattacc gccaaatcct tttgatatcc tttttcccaa 6840
atgtccgcaa ttacacacat accctcgcac tagcgcccct cgtttttca gggtataaat 6900
gtatggggaa aacacttgta aatagcggca gcagccac 6938

<210> 182
<211> 2010
<212> DNA
<213> Drosophila

<400> 182
atggtcccaa gtcgggggaa gtggccgtgc cgcatgtttt acgactacac cgacatcgcc 60
gaggatgccg cacgccagtt tattgaattt ttatcgggca aattcccaaa tgccaacaca 120
ccgattgcca tcgatgagcc gacacgtgcg gaggtgagtc gccgggccaa tggaatcgcc 180
agctacgccc tcaacgagga cgacaatctg ctggcattcg ccattgccgc gcccagcatt 240
cacacggtgg tcgttaaatt cagggataac gttacgatac caccggatca ggtgcacaat 300
aaggcatatt tgggttccta ttggcgggag ttgggtgcgg cctggaacag cacggacggc 360
acccaggagt ggggcgctcc atttcgcgac tgcaacctgt tgacgcgtcg ctggctctgg 420
ccatttcgca tatccttcag cgagcacagg atcaaagttg tggcggctgc cttcattgcc 480
gccgacgagg atgtgtgcaa cgatggcctg gaggaagttt tcggtcgccg tcatggttgc 540
gatcggaata cgaccttctg cctgctcacc gagaacaaac ccgccgccac ccgggatgtg 600
tacacctgtc tgtgccggga atcctactac ctgcccaatt ccacgctcca gggattccgt 660
ggcgatcgag tggaactgtc cgagggctac gacaactact cgtgcattcc gtgtcccggc 720
ggatgcacta actgcgatag caacggcgtc tgtctgacct ttcaggagga ggaggtgctc 780
aatgtggacg cctgtctgcg cctcctggtg gccatcgtac tgggcgcctg catcctgtgc 840
tgcattgtcc tcggggtgat tgtcttccgg cagcgaaagt gcaaggccat tgcgtctggt 900
atgtggactg tgctggagac gatactgctg ggcattgttt tactttatgc atctgttgcc 960
gtccatttct ttcccgcatc caccgagcgc tgccttctgg agccctggct tcgggagctg 1020
ggcttcatca cctgctacgg cgccatcata ctgaagttgt accgccacct ggtggacttt 1080
cgcacccgga aagcgcatcg ctgggtgctg cgcgacgtgg acctgctcaa gtatctgggc 1140
accatggtct ttgccgtcat ctgctacatg gccgccttta cggcctcgtc gttggatctt 1200
ctcgaaagtg cgcagctgga gagcctcagg gaggcggaca cgaacacctg ccatccgctc 1260
aagtgggaac tggtcacgca gaccagcgag atgctcatcc tgtgctttgg actgcacctg 1320
tccatcgcca gtcggaatgc caacacccag ttccgggaac gacaatttct ggtcaccgca 1380
ctgacgctgg agttcctggt ctcgtccagc ttctactttc tgcgctttgt ctacctgccg 1440
gaaatgagtc ccagcgccat cttgctggcc ctgttcatcc gctcccagct gacgaatagt 1500
ttcgccttgg gtctgatatt tgtgccaaag ttgtggtacc agcacaagca gggtacgtca 1560
cacgacgccg gccagcggct gggcggagga tatgccgggc tctgtctggg cgatccggac 1620
attggtgagc tgaccatatc cgaaatgagt cccgaggaca tacgcgccga actcaaaaga 1680
ctgtataccc aattggagat aatgaagaac aagactctgc ggcaggacaa tccgcacatc 1740
agtaagcgac gtggcggacg caaggcgggt caccgtcgct tctccctgca gaaaaagggc 1800
agcaaggata aggctctaag tgccaaacac cgcagcaaca agcatcatca ggacatcgag 1860
atcaccgagg cggaaccttc caggacgccc gaggattcgg tgtgcagtgc cgaaggaccc 1920
acggatacct atgcggaaat ctccggcgtg tcccactcaa tgctctctca ctcgatggtc 1980
tcccactccg tggtctcgca ttcgaagtag 2010

<210> 183
<211> 669
<212> PRT
<213> Drosophila

<400> 183
Met Val Pro Ser Arg Gly Lys Trp Pro Cys Arg Met Phe Tyr Asp Tyr
1 5 10 15
Thr Asp Ile Ala Glu Asp Ala Ala Arg Gln Phe Ile Glu Phe Leu Ser
20 25 30
Gly Lys Phe Pro Asn Ala Asn Thr Pro Ile Ala Ile Asp Glu Pro Thr

```
        35                  40                  45
Arg Ala Glu Val Ser Arg Arg Ala Asn Gly Ile Ala Ser Tyr Ala Leu
    50                  55                  60
Asn Glu Asp Asp Asn Leu Leu Ala Phe Ala Ile Ala Ala Pro Ser Ile
65                  70                  75                  80
His Thr Val Val Val Lys Phe Arg Asp Asn Val Thr Ile Pro Pro Asp
                85                  90                  95
Gln Val His Asn Lys Ala Tyr Leu Gly Ser Tyr Trp Arg Glu Leu Gly
            100                 105                 110
Ala Ala Trp Asn Ser Thr Asp Gly Thr Gln Glu Trp Gly Ala Pro Phe
        115                 120                 125
Arg Asp Cys Asn Leu Leu Thr Arg Arg Trp Leu Trp Pro Phe Arg Ile
    130                 135                 140
Ser Phe Ser Glu His Arg Ile Lys Val Val Ala Ala Ala Phe Ile Ala
145                 150                 155                 160
Ala Asp Glu Asp Val Cys Asn Asp Gly Leu Glu Glu Val Phe Gly Arg
                165                 170                 175
Arg His Gly Cys Asp Arg Asn Thr Thr Phe Cys Leu Leu Thr Glu Asn
            180                 185                 190
Lys Pro Ala Ala Thr Arg Asp Val Tyr Thr Cys Leu Cys Arg Glu Ser
        195                 200                 205
Tyr Tyr Leu Pro Asn Ser Thr Leu Gln Gly Phe Arg Gly Asp Arg Val
    210                 215                 220
Glu Leu Ser Glu Gly Tyr Asp Asn Tyr Ser Cys Ile Pro Cys Pro Gly
225                 230                 235                 240
Gly Cys Thr Asn Cys Asp Ser Asn Gly Val Cys Leu Thr Phe Gln Glu
                245                 250                 255
Glu Glu Val Leu Asn Val Asp Ala Cys Leu Arg Leu Leu Val Ala Ile
            260                 265                 270
Val Leu Gly Ala Cys Ile Leu Cys Cys Ile Val Leu Gly Val Ile Val
        275                 280                 285
Phe Arg Gln Arg Lys Cys Lys Ala Ile Ala Ser Gly Met Trp Thr Val
    290                 295                 300
Leu Glu Thr Ile Leu Leu Gly Ile Val Leu Leu Tyr Ala Ser Val Ala
305                 310                 315                 320
Val His Phe Phe Pro Ala Ser Thr Glu Arg Cys Leu Leu Glu Pro Trp
                325                 330                 335
Leu Arg Glu Leu Gly Phe Ile Thr Cys Tyr Gly Ala Ile Ile Leu Lys
            340                 345                 350
Leu Tyr Arg His Leu Val Asp Phe Arg Thr Arg Lys Ala His Arg Trp
        355                 360                 365
Val Leu Arg Asp Val Asp Leu Leu Lys Tyr Leu Gly Thr Met Val Phe
    370                 375                 380
Ala Val Ile Cys Tyr Met Ala Ala Phe Thr Ala Ser Ser Leu Asp Leu
385                 390                 395                 400
Leu Glu Ser Ala Gln Leu Glu Ser Leu Arg Glu Ala Asp Thr Asn Thr
                405                 410                 415
Cys His Pro Leu Lys Trp Glu Leu Val Thr Gln Thr Ser Glu Met Leu
            420                 425                 430
Ile Leu Cys Phe Gly Leu His Leu Ser Ile Ala Ser Arg Asn Ala Asn
        435                 440                 445
Thr Gln Phe Arg Glu Arg Gln Phe Leu Val Thr Ala Leu Thr Leu Glu
    450                 455                 460
Phe Leu Val Ser Ser Ser Phe Tyr Phe Leu Arg Phe Val Tyr Leu Pro
465                 470                 475                 480
Glu Met Ser Pro Ser Ala Ile Leu Leu Ala Leu Phe Ile Arg Ser Gln
                485                 490                 495
Leu Thr Asn Ser Phe Ala Leu Gly Leu Ile Phe Val Pro Lys Leu Trp
            500                 505                 510
Tyr Gln His Lys Gln Gly Thr Ser His Asp Ala Gly Gln Arg Leu Gly
```

```
        515                 520                 525
Gly Gly Tyr Ala Gly Leu Cys Leu Gly Asp Pro Asp Ile Gly Glu Leu
    530                 535                 540
Thr Ile Ser Glu Met Ser Pro Glu Asp Ile Arg Ala Glu Leu Lys Arg
545                 550                 555                 560
Leu Tyr Thr Gln Leu Glu Ile Met Lys Asn Lys Thr Leu Arg Gln Asp
                565                 570                 575
Asn Pro His Ile Ser Lys Arg Arg Gly Gly Arg Lys Ala Gly His Arg
            580                 585                 590
Arg Phe Ser Leu Gln Lys Lys Gly Ser Lys Asp Lys Ala Leu Ser Ala
        595                 600                 605
Lys His Arg Ser Asn Lys His His Gln Asp Ile Glu Ile Thr Glu Ala
    610                 615                 620
Glu Pro Ser Arg Thr Pro Glu Asp Ser Val Cys Ser Ala Glu Gly Pro
625                 630                 635                 640
Thr Asp Thr Tyr Ala Glu Ile Ser Gly Val Ser His Ser Met Leu Ser
                645                 650                 655
His Ser Met Val Ser His Ser Val Val Ser His Ser Lys
            660                 665
```

```
<210> 184
<211> 4243
<212> DNA
<213> Drosophila

<400> 184
ctcgtgacct ttgaacccga cgctggcact ttgaaacgga aacggaaaca cacactcaca 60
cacacgtaat ttggttagtc tgatatcgat gcactatatg ttggctaatt tagacggact 120
ttggttttga ccaatcgcca gggtcctaca tatatggttt tgactatata ttctcgctat 180
gtttagttga ttatacgatt taggggttta ctttaggtat tttgagcatt ttcaattcat 240
gcgattgtcg ttgaaataac aattgggttt tagctttgct ttagctgccg tgacttttgt 300
ttgtatctct ataaatttat caatattctc cgattctaat tgcgttgggg gtgtggttgt 360
ttattcgcag gacttctcac gtagttgtct ataatttgcg gtctcttaat tggcacagtt 420
ttagttttag ctttagctcg tcgattttgc gaatccctat gtacacgcat agctcagtga 480
gttgcagtga ttgccatttt gaataattta tttgtggtta tttcttttcg tgagcaaaat 540
gctttacggc ttacagattt tcgattggcg tacgtacata catattacat attcatatac 600
atttatcgaa atgagtttac aacctgtgct taatttagtc gtacaaaaat actcagtatt 660
tacaacaaga tataaacgac ataaaatcat aaataaataa gttatcctat acaaagtccc 720
gatctgtgct ccaaggaacc cggcaaactc gaattgagat taaatgtttt aaaaacgaaa 780
acaaaaatga aaaaatgtga agaaaatgaa atggcacgct ccacagattg aataatgtca 840
acattacgag taaatcagcc gacgcatttg ccggcatatt aattttacaa attaatggaa 900
cactttaatt gcattaaatt gattccaccc atccggctgg catcgattga ttcgattcgc 960
atccttacac cagtcgcacc atttgcacca tttgcaccat ttacacatga ctcgaggccg 1020
tcgaggcggc ggcattgttg ttcttgttcc atttgttgat gtactccatg ggtatttcaa 1080
tctggatctc ctgcacatcg tcgcccgcca agcaatcgac attggatttg cccttgcccg 1140
catccttgcc cgtatccttt gcatccttcg actccctgcc atccttggcg gcagctgcca 1200
attgtccgcc caggcgggat ccgcccgccg ccagccgctg ctgctgctgg gcgatggggg 1260
tcagcagcgc ctcatccctg tattataagt tgtccaccat ctgagactgg cggcgatgcg 1320
agggcgttgt cgagttaagg tagctctccg agtagacgct cttcgcccgc tggtcgtagc 1380
gcaggctgga gcgccggatg tcgatgtcca gaggcaggtg gcgatccttc caccaattgc 1440
agaacaggca gagcagcgtg ttgcggaacg cctcccggaa gtcgcggttg aagtacgcgt 1500
agatcagcgg gtttagcgtt gagttgaagt acccgatcca gaagaggatt gagacgacta 1560
tgtccggcac ttggcactcc tcgcaggtca tggagagtgt gtacctaaaa ttaaaacaat 1620
ataattatta gcttacactc attatcttta acttaattta agtcaatttt aaattagttg 1680
cacaggtaat gcaattttat gtactttttt tttcccggca ttgcataata tctatttttt 1740
aataatacaa aagcaaatat aaaagatcaa atgaatattt ttgtttcata aaagttaaac 1800
ctcataaacg atctatattt gaatactttg atcaccattt gaaatctttc tgcatcactt 1860
tttgtctgtg tagcacataa agatgagata tatataccca taaattccag atgcatgagt 1920
gggccagggc tttgcataaa tgagattata ctctgggtgg tgtgaaaaac aacaaccgcc 1980
```

gaaacaatgg cgcacgggcc gtatgagtaa tatttaatta tgcatgcttg acattagata 2040
aattatgcct ggaatcccat gtcagcggga ttggcctcca actctcaacc aaacgttcaa 2100
ccttcgacgg ctggttggta ctcaccacag gaagaaaggc agccagcaga ggatgaaggt 2160
gcccatgatg atgcccagcg tgcgtgcggc cttgtgctcc cgcttcattt tgattaagtg 2220
cttgtcctta gtgggggtgt gctcctgctc gacctcgtgc agcgtcaatg ttttcgacga 2280
cccggatccg ctcagcgcct cgcctgccgt aggagataag atggagattt gataaaagtc 2340
caggctggag atgggccact gaactcaatg cactcgcact gcatcactgt ccaggatgct 2400
tgtagaacaa actagtgatc tttcacgctc actacctgtt tgctattcag aaagtaataa 2460
gactttctgt cctggtaaat ttatattttt gaagtattat ttccacttaa aacccattat 2520
atatattcaa acaaacattt ctttacagcc aaactaggta agccttgttg tttacaaccc 2580
aactttgact ccaaactcac ctgatggctg catggatggt cggtgcatca gcatggcatt 2640
gccgtgccgc atcatcagct gcttctcctg ccgattggct tcccggaaga tggccaggta 2700
ggtgaatatc ataatggtgc agggtatcca gaacgatatg gagctcgaga tgacggcgta 2760
gtacttgttc accacgaacg agcactgcgt cggattctgg atgacgaact gctggtgctg 2820
cggcgtggtg taccagccga tgaagatggg caagaaggag agcagtgccg gcgatatcca 2880
tgtgtttagc agcataatgc cgaccacgcg tttcgtcatg ctaatcggat acttgagggg 2940
cttaacaata gcatagtatc tgtgaagtgg gcaaaacgcc aggttattgg gccctcatat 3000
agctatagtt ctggctgtgg catgcaactt acctatccac agatatgcag cataaatgca 3060
aaatactcgc tgttgagaag tagacatcga ggctgttcca caaatcgcac aggaagggc 3120
tgaagttcca cctgaaaggg tccaatgaat ggaggcatta gtatcgcgat gcttacccaa 3180
taacgaggca ataaactttg gcctttactt tctaacgtca ttgtatctga ttatcgagtc 3240
catgaatttg cagtatatat ttttgttgac tcatttctgg ttttttgcct cacctaaaat 3300
gccacgcgaa atactttat gaagtcatcg ctttaaaact gatttcctgg tttttagctg 3360
tgcaaaaaat ttagtttttc acgattaatt ccctttgtga caacacagtt atttacgtac 3420
tctgattaaa atcatttgat tgtgggtaaa attgcgctag ttctaatatt tactaccatt 3480
taaaatacag atatataggc cataaaaact cactagatat gcgcaataag gttgctctaa 3540
atttaaacgg cattttactg ccaattcctg aatgtaggtc tataaaaaac tcaaactgac 3600
tagacacact ctctttttcg cctattttcg ttatcaggcc gcctgattta tgtgtttatc 3660
gatttcatct tagtttcggt tgccaggtga cattgcccct gaactgccat tctggagggc 3720
caaattaaaa gcttagatgc actgtgagac cagaaaatag gccataaggc ccgaggtgaa 3780
aagtggcaag gaaagcggag gcgactctgt tcgctgcatg aaaaactcat taaaacttgc 3840
gtcgtcgctg ttcgcctgtc gcttctgccg tcgcccttgg ttgccactcc gcattttggc 3900
catttcggtg aaggcggata cccggccttt ttggggccag gagccagtgg aatttattat 3960
cgggcgcccc gtttggcatc agtttcaaaa acttttcgac tatccaacga ggctgttttt 4020
gctcgctggt ttttggtgct ttacactaca ttgccatttg gggcgggcgc gccatcggtt 4080
ttgacactta agccatgaaa tggcagttcc tcgtgctgca attatgcaac tggcaactgt 4140
gaaggtaaaa tgtctaatga atgccaagtg gcggcgagtt ggctgttgcc agaaccaaac 4200
cgatgccaat ttatatgcaa ctcgcgcttg gccaaagttt ttc 4243

<210> 185
<211> 1224
<212> DNA
<213> Drosophila

<400> 185
atggactcga taatcagata caatgacgtt agaaagtgga acttcagccc cttcctgtgc 60
gatttgtgga acagcctcga tgtctacttc tcaacagcga gtattttgca tttatgctgc 120
atatctgtgg atagatacta tgctattgtt aagcccctca agtatccgat tagcatgacg 180
aaacgcgtgg tcggcattat gctgctaaac acatggatat cgccggcact gctctccttc 240
ttgcccatct tcatcggctg gtacaccacg ccgcagcacc agcagttcgt catccagaat 300
ccgacgcagt gctcgttcgt ggtgaacaag tactacgccg tcatctcgag ctccatatcg 360
ttctggatac cctgcaccat tatgatattc acctacctgg ccatcttccg ggaagccaat 420
cggcaggaga agcagctgat gatgcggcac ggcaatgcca tgctgatgca ccgaccatcc 480
atgcagccat caggcgaggc gctgagcgga tccgggtcgt cgaaaacatt gacgctgcac 540
gaggtcgagc aggagcacac ccccactaag gacaagcact taatcaaaat gaagcgggag 600
cacaaggccg cacgcacgct gggcatcatc atgggcacct tcatcctctg ctggctgcct 660
ttcttcctgt ggtacacact ctccatgacc tgcgaggagt gccaagtgcc ggacatagtc 720
gtctcaatcc tcttctggat cgggtacttc aactcaacgc taaacccgct gatctacgcg 780
tacttcaacc gcgacttccg ggaggcgttc cgcaacacgc tgctctgcct gttctgcaat 840
tggtggaagg atcgccacct gcctctggac atcgacatcc ggcgctccag cctgcgctac 900

gaccagcggg cgaagagcgt ctactcggag agctacctta actcgacaac gccctcgcat 960
cgccgccagt ctcagatgca gcagcagcgg ctggcggcgg gcggatcccg cctgggcgga 1020
caattggcag ctgccgccaa ggatggcagg gagtcgaagg atgcaaagga tacgggcaag 1080
gatgcgggca agggcaaatc caatgtcgat tgcttggcgg gcgacgatgt gcaggagatc 1140
cagattgaaa tacccatgga gtacatcaac aaatggaaca agaacaacaa tgccgccgcc 1200
tcgacggcct cgagtcatgt gtaa 1224

<210> 186
<211> 407
<212> PRT
<213> Drosophila

<400> 186
Met Asp Ser Ile Ile Arg Tyr Asn Asp Val Arg Lys Trp Asn Phe Ser
1 5 10 15
Pro Phe Leu Cys Asp Leu Trp Asn Ser Leu Asp Val Tyr Phe Ser Thr
20 25 30
Ala Ser Ile Leu His Leu Cys Cys Ile Ser Val Asp Arg Tyr Tyr Ala
35 40 45
Ile Val Lys Pro Leu Lys Tyr Pro Ile Ser Met Thr Lys Arg Val Val
50 55 60
Gly Ile Met Leu Leu Asn Thr Trp Ile Ser Pro Ala Leu Leu Ser Phe
65 70 75 80
Leu Pro Ile Phe Ile Gly Trp Tyr Thr Thr Pro Gln His Gln Gln Phe
85 90 95
Val Ile Gln Asn Pro Thr Gln Cys Ser Phe Val Val Asn Lys Tyr Tyr
100 105 110
Ala Val Ile Ser Ser Ser Ile Ser Phe Trp Ile Pro Cys Thr Ile Met
115 120 125
Ile Phe Thr Tyr Leu Ala Ile Phe Arg Glu Ala Asn Arg Gln Glu Lys
130 135 140
Gln Leu Met Met Arg His Gly Asn Ala Met Leu Met His Arg Pro Ser
145 150 155 160
Met Gln Pro Ser Gly Glu Ala Leu Ser Gly Ser Gly Ser Ser Lys Thr
165 170 175
Leu Thr Leu His Glu Val Glu Gln Glu His Thr Pro Thr Lys Asp Lys
180 185 190
His Leu Ile Lys Met Lys Arg Glu His Lys Ala Ala Arg Thr Leu Gly
195 200 205
Ile Ile Met Gly Thr Phe Ile Leu Cys Trp Leu Pro Phe Phe Leu Trp
210 215 220
Tyr Thr Leu Ser Met Thr Cys Glu Glu Cys Gln Val Pro Asp Ile Val
225 230 235 240
Val Ser Ile Leu Phe Trp Ile Gly Tyr Phe Asn Ser Thr Leu Asn Pro
245 250 255
Leu Ile Tyr Ala Tyr Phe Asn Arg Asp Phe Arg Glu Ala Phe Arg Asn
260 265 270
Thr Leu Leu Cys Leu Phe Cys Asn Trp Trp Lys Asp Arg His Leu Pro
275 280 285
Leu Asp Ile Asp Ile Arg Arg Ser Ser Leu Arg Tyr Asp Gln Arg Ala
290 295 300
Lys Ser Val Tyr Ser Glu Ser Tyr Leu Asn Ser Thr Thr Pro Ser His
305 310 315 320
Arg Arg Gln Ser Gln Met Gln Gln Gln Arg Leu Ala Ala Gly Gly Ser
325 330 335
Arg Leu Gly Gly Gln Leu Ala Ala Ala Ala Lys Asp Gly Arg Glu Ser
340 345 350
Lys Asp Ala Lys Asp Thr Gly Lys Asp Ala Gly Lys Gly Lys Ser Asn
355 360 365
Val Asp Cys Leu Ala Gly Asp Asp Val Gln Glu Ile Gln Ile Glu Ile

```
    370                 375                 380
Pro Met Glu Tyr Ile Asn Lys Trp Asn Lys Asn Asn Asn Ala Ala Ala
385                 390                 395                 400
Ser Thr Ala Ser Ser His Val
                405
```

<210> 187
<211> 11847
<212> DNA
<213> Drosophila

<400> 187

```
gtgaatcgtt tttgcagcga atggcgaaat ttcctacagc ttgccaattg ttttggttgt 60
tcgcagttca ttctttttac tcattggtaa gattgtcttt ctgcatgaaa attcataacg 120
tcagctacac aattctcttc cattcaatat tttgtccata cataatttcg attgcgtaat 180
ttgtaatttt caataactct gttatgtaga tttggtttgc tagttttcta ggttacatat 240
gtatgatact tagagctaga ggtaacatta tgggattgca ctttttccat gatttaggct 300
agttctatct gtaacaacat taaaaagtgt gtttaatatg aaataaaagc agggagttca 360
ttccgctcag gatttgcaat cctaatatac gatgatgaca catgtgaata atccaatttt 420
ggatcgtatt aatactccaa aaatcgatgg aaaaagtgtt tcgccttcga caagccgaag 480
tggtgtacat aattcttaga tgattcattt gacaacattt cacttgaagt tggaactttt 540
ctacgcaaag agtcgctcga tttctaattg gctttcgctc tgccatatat gtatagttat 600
tacgtatatc ttctgtttag ttggacaaac aacacttata tatggatagt atagttccgg 660
ggcagcattg tccctgcagc cgcggcttgg cgggcttaga acacaacttt aatttaaatt 720
ccctatcgtt gcttaaacga aattgtacac gtatacataa acacacatcg agtattgctt 780
agcttcaagg attatgctag ctatatcctg gcacaaatgc cactggtatg gccgttgctc 840
ttgttacttc tcttcacttc agctggtgtt ttccctcctt gggcacatat acatgtgcga 900
cgattcactc aagtattaaa tatcacttat aggggcattg ttccgcatcc acgagggtgg 960
acgaagcagg aatgtggccg ggatgcagac gctatccggt ctagacgtag taggcgttgt 1020
acggatttcg tttctccaga cagggcagca cacactccag cgtattgcgg aatgcctcgc 1080
gaaaatcgcg attaaagtag gcgtatataa gcggatttag cgtggagttg aagtaaccga 1140
tccagaataa caccaccacg agcacatcgg gacatgggca ggccggaccg cagagcgatg 1200
tgatgacata cctggaagga ggagtgagaa tcacatgaat atgagtcacc agaaaaagaa 1260
gaataagaag gccaacaatg gattacggat ggcggtacca catccgtacg gcctatctac 1320
cgtcactcac cacagaaaga agggcagcca gcagagcaga aagacgccca tgatgatgcc 1380
cagggtgcgg gcggccttgt gttcggcctt ccatcctttg gccggtcgca cggaaccgcc 1440
tgtggggcag atagagagtt tagaagggat atgaaagctg gcaattggct gattaagttc 1500
aattcgaatc gattcgtttc gatccgccag tctgtttttt ttcttttgtt gggtcatagc 1560
aatcgacagt tttctggttc atttctgtct gtctgcgaaa gcactcgaaa aatgggtttc 1620
tttttaagta gtctataata tgcttaatgt ggtatttgtt gtctacctgt tggatgtgtt 1680
tgttttatgg tttattgctt ctaggatgaa attgtttata cagctaatac tggaaagtat 1740
aggtgtttgc ttaagcttta atcggcgata cccttttatg ttaactatgt atgtaaatta 1800
attgaacttc aatcaattaa aatacagaaa atttcaaaca aacaaatttt aaagaggtaa 1860
ctaaactaat cttaaaaatt attaaatatt ttctctgtgt gtaaattgcc aagaaatgcg 1920
gtttatgttc gttttattgt gagagaaaga ttatgccggc atgttgcgta tacgccatgt 1980
gtgtggtatc cttgtttaac tttgacatta gtaatacttt tgctgttata tgttgttggc 2040
tatcgatggc atttcaaacc ccacaccgcc ccgctccctg cgtccacctt gcaagtgact 2100
gtcattaaag tgctggaaag gccaacatgt gtgcagtatc ttggtgctcg ggcgtgtggg 2160
tgtgcgggaa aagcaataaa aacgaggcaa acaattgaaa atatttacac aggacacttg 2220
ttatcacaaa gacatacaat cccccgcatc catccagtcc agtgcgcccg ccccaaacat 2280
aaacataaac atcactggcg gccaggagga aaaacaaatc agacagacaa cagtgagcaa 2340
caggaagcaa cggaccatcg cccttgttca cccattgccc atcgcccatt gcccatgttc 2400
ccccatccag tttgccacct ggatcccttg atggacgcca ctcggtgatg gccacagcca 2460
ggtgaaatca taatgcaaag gctgtgtgga gggaggcgta ataaagcaaa gcaaacgcac 2520
gtgcaaacgg aagcaattat gtacagcatg agagcagcga aaggtgaagg attcggatac 2580
ggattcggag gcagattagg tggaggtgga gacagcatga ggtttcgatg gtgtaacact 2640
ccggattcag ttgcctaagc tggttagctt ggcatccgga taaggaggtg ccaaaaaact 2700
ggcacttctg tgggtatttg ccactaaagc tgggggctac aaaaatttga ggtagttctt 2760
ttcggtttca tttcacggtt gatttaatga tacttaggtt aaccaccgtt gaaatgattg 2820
```

ggattgcaga aattcaggga tatctattct agtatgattt ttggtaagaa cgttaactag 2880
tttcaggtgt attcaaaagc tttcaatttt aaacataata ggtttaaaca tatacggaac 2940
agtaatattg attgtttttt gactgacagg atatttcata tgaaataaat agattttga 3000
tatgtttggt ctaatgcagt ggcataaaat acgaagtgaa tgggacttga tttgattaag 3060
atttctgat gattatatat gtattgatat agagaaacta taactattgg tgttgtgaaa 3120
ttagtgaatt ctgagtgtga aatgtattat aatataatgc acaagttctg gactgtggga 3180
aatcgaatga ttcagatgga taatgggtag caatgagata ggtgctgaag ctgagacaag 3240
cagaaagaca gagatagaga gagaagcggg agagaaagag cgaaacgtag agactggggg 3300
aggagaggag gggagatacg acaagactca ccggttgcac cacagtgact gcatccggtt 3360
ttgaagcggc gcttccggtt cttgctctgc tgccaacatg ctgtgaccgc cgccgttcgc 3420
ttgtgccaga atcgagtgga actggacgag gaggcgggcc tctggccaag atctggagtt 3480
tgctcgtgaa tactacgttt actactagtg gttactgtgg tgtcgattgt gctgtcggcg 3540
ttggcggttg cggttgcgtt tgcggtggtg ccggtggtgc tgctgatgtc ggcgccgcag 3600
agcggacact tgggcagcgg actgggcacc gcggcgaagg tgatgccgcg gggccgctgc 3660
tcaatgatct ccgcctcctc ctcatccaca aagatgcaga tgccgccgcg tcggatcagc 3720
gaggatttat tggaaaattc gtgctctgct tgcagttgat tgttgcattt tcgtacgatt 3780
tcgattgtca atttggccat ggcattttg gggttttatt ggttggattt ggttttggag 3840
tagaagtagt ttgtttgttt ggttggacat aatttggtga tttcgattta catatataca 3900
catacaacac acaacgaaat ggcgtggaag agagatagat ataaagagat agagacaagt 3960
agtaaacaat gtacacaggg tatttcacag gatacgaaag gacttcaact tctgagatca 4020
catgcagcca tgcaccggca aacgcaaagg atcgggatcg aaaagggaat ggattcagat 4080
tcattatcgt tggctccctg tctctggaag taggtcgtcc ctttctgtct tcgtttgcag 4140
gtgcaactaa ctaaatgcta tacattatat tatatatata tagtaagcta gttaactaaa 4200
gtttcattcg aagcatttga gtattacttt tccaatgagg gtttataggt tctatacatt 4260
ttaaaacact gtatgatttt tcagaaattc gcaccgcttt tgaaaggaac ttaaatacac 4320
aattcacagt ttcttcaag gtaagaacta tatttaaata tggtaaaagt tcgtatcttt 4380
attttaagat caaaattcta gtaagaccag attccttatg ttaaacccga attcaggcca 4440
attactcctc aaaaaaacaa ggacaaaaga atgtgcgaaa aagttaagtc aaacattgaa 4500
acacttgaac ttttcgtgga aaactgacaa aaacagaacg aatttctctg ggcagatat 4560
aaattgtctt aaaagagcta cttaaatcaa gcacgagcta tacagtctac acaatgacca 4620
aatggatgac tactcacttg cggaactggt ggaaacgcgg cccatcggga tgtagctgcg 4680
actgtcgctg ttctcctgcg agaagttgat gacattcacg tcgtaggtca ccacggacgc 4740
ctgtctcttc cgcacagttt ggatgccatc gtagcctgcc aggatgctgc agcggccttg 4800
gcggccaccc actcctccaa cgccaccagc ccctccattg ccggaaacgc tcacgtccgc 4860
atcgccggcg ctatctttgc gcttgaacag tggcgacgtg acatcccggt gctcatcgcc 4920
gcatccctca accggcgaca gcgtggaatc gagtatgatc tccggcttgg gcgatatgtt 4980
ggactgtgag gccatttggg gatttggctg cggcactgga acgggcatag gcaggctggt 5040
caccacatcc aaggacacag tggatgtgtc actgggcgta taaccaggtg cggcggcatg 5100
cggcggcggc gataggaagt acacattgat ggacgccgac tgcgcgttgg aagaggcgga 5160
agagtccatg gcacctccgt tcggcacgga aatggccggc ggagccggtg tgttcggtgg 5220
ggacatcaga tagttggcgc acacctggtc actgagcgtg ctcaggcgct gtggagcacc 5280
agccggtcgg cccagcggaa agctcagcga gcgaaagatc tcggccagtt ccgagctgct 5340
ggcaatgagc aggccactcg gttctggcaa ttctggcggc gataacgatg gtggcggcgt 5400
ggccggctgc tggggcgact ccccaaagca atcggccagc tgcgccaact cgtcgtcgct 5460
ccgaatcctc tcgggcggtg ttttcagctc actgaagtca atcagctgta ccccgtcgct 5520
gaactgctgg atcacactgg ctccactctg ctcgctgagc acattgatgt ccagcaggat 5580
gtccggcttg ttgggcgtgg agtccgtatc gctgtgctgc cgctggcgtc tctcaagttt 5640
gcgttccttg ccgccaccgc atccgttgta gctgtgcgac aatttccggt tccgggagcg 5700
ccggaagaag ctggatccct cggagttggc ctgatgtcgc ttcttttcca gcgtaaagca 5760
ctcggtgctc gagcgtttcg tcttcgacag cagggacaga atgttgaagc ccttctcctt 5820
cccgggcact ggctgctccg gcaacgcccg cgattcacta atggtgcgca ggaactccga 5880
atcgctggag gcgaatgcct tacgcccaca ggccgctgcg cccgccttca gcctgatgaa 5940
gtccgtttca ctgagggcag cgtactttcc tccgcctggg ccactgataa tggtgccgcc 6000
cgtctggcgc ttaaagtaca gatagttgtc ctctatcagg cgcttcagct ccttgttgtc 6060
ctcggccagc ggcaggtggg tgtcgtacag gtcgttgccc tcgctctcgt cgctgttgag 6120
accaccaaag acccggtacg cagtggactc cggttccggt tcgctgttcc cactgccgct 6180
ggtcaggctg agctgctgtt gcgggccatt gggatgagaa tgctccttat cgagctcctt 6240
cgaagggatc atcttgccct ccgtccatat ctgcggattg aacactggcg ccggtggctg 6300
caactgattg tgggagggct gctgttgctg cagtgccatc atcgagggag cactgtccgt 6360
gtgggtgacc ttctcgtacc gctcctgctc caagtccctc agatcgatgc tgctgcggct 6420

gaggcggcgt ggcggcggcg atggcactct cagttcatcg cactcgtccc tatcgtcgtc 6480
ctcgtcggtg gccggctgca gcatgtcctg tgggcaaagg acaatcactc aaagggaagt 6540
catggtggcg catcaattac gcacagccac aagggaaaag gaaaaactct caagcggagt 6600
gcttccctct agcacctaat ctaatcgctc cacatccgtt cattgaggat gggcaaaaag 6660
ctttagagca ggatgtcaat gttgcccgtt tgtatacatg aactgcattg ctttaaagtc 6720
tttgtatatt tatttaggat actatctgga tgtatcagca actttaccaa attttactca 6780
ataaaaaata aataatgttt tattcaacta ctaagagatc ttcagaatgg caagcaaatg 6840
cgagtactaa agcatttatc ggtttaataa aaataaaaca tttttaagta tctctaggta 6900
ttttaggcat ctttggatta atgtgtattt tttcattttg aacatctttg aaatgtttat 6960
tgtttattga atgatttatt gaatttatat caataaaatg cagcaaaata ctcttttggc 7020
tttcaataaa tgacatctgt atcttttcgg ttcagtaaca tttagtaacc ttccaaactc 7080
acctccaagt tactaagcgc actgtgcgtc tcctcccggg cggatgtggc attcagatca 7140
cagtcactgg gatgcagata gctcaggctg gtgggctgct gggtgtggcc catgtggacg 7200
ctgttaagga ggatgttgga gctggtgcgg ctgagggcct tgcgttgacg aatcgcctcc 7260
ctggaaaacc acagagcaca gatcagcaat taagcaaact catccactac gaccagcgac 7320
catcgaccat cgaccatccg ggtaagaact cacttaaaga tgcgccagta catcaccagc 7380
atcacaatgc cgggtatcca gaagctcacc gaactggaga tgagggcgta ggccttgttg 7440
accacaaacg agcactggtc cggatgcagg gaaatctccc gcaggtgctc ctccgtcgtg 7500
taccagccca gaaagatggg cgtgaaggat atgagggcgg gcaggatcca cacattggcg 7560
agcatgaagc agaccgtttt gtgtgtcata ttcaatggat actccagtgg acgcacaatg 7620
gcgtagtatc tgttaaaaag aagagggggt catctttatt aatggaaatg aattgcggac 7680
aggtcctccg cacaaaggtg ccaggtggct ttcagccatt tcctccttct cctcctcctc 7740
cttttggctt actaattgat tatcgctctt tggttggctg aataattgtg tttcggtact 7800
caggttgcca tgcacagaaa tcgatttatt aacctttttaa agatcatgtt gcatttaaat 7860
gaaggataaa tgctcttaac cgaattcatt atcggaatcc gtgaaaataa ggtggtgtat 7920
tctttttaat tatcaatata ttaataaatg caattgaaat gaaataatta ttaatccacg 7980
tatattgtac ctcattattt tatcgaaact taatgctcct tttaaagagt aatgattatg 8040
tttggaacat tacagattta ttagtttagc caaatttccg ttagtggcca tcattaattc 8100
tctaattttg ataggttctg gttgaatttc cctttgtttc ttggtatgtt ttgtggcccg 8160
tcgaaaggtg agtgaaacac aaattaatta aaacagacaa aaggcccatg actccagtgt 8220
ccatcgacca tccctccatt gttgctggtc agcccgaaaa tcgcagctgt tggtgccttc 8280
caagtgggga atggcgccaa ttaacggctt gacaggatgg agaaggttgt gggagttggg 8340
tggaatatgg gtggaatatg ggcaacgccc acgtttcggc cagctgcaaa tcatttgggc 8400
cgcattcatc atacgggcgg gcctaatgtt tgcttttgat gtgcgcgcct ttcatagtaa 8460
ttaattgttt cttcttggca attcgccggg ctttggtatt gaaagtgatt tggttgccag 8520
gcatctgggc aagattaatg gtatttggcc aaactaaacg gctttcttca gtaccatagc 8580
ccagcacaat tgactttaaa ccgcttaata tatagataga gagcgacata aatgggcact 8640
cacctgtcca ctgatatgca gcacaggtgc aatatgctgg ccgtggaaaa gtaaacatcc 8700
aggctgttgt acacgttgca catgaacggt ccgaacatcc actttccgcc ggacagttcc 8760
acggaggcgt tgaatgtcat cgcacagagg gccaccagca tgtcggccat cgccagcgac 8820
acaacaaagt aattggttat caccctgcga acgaaacaaa agcccaggaa tttgtgaata 8880
tggcgtggtt ctgggcggga tttgaccagc ttctccccac atctaaccac ttgtcacagc 8940
gcaaaaccac gccaaacgta atggcatatg gcatgcatca tatgtgcaca tgccgctgta 9000
ctaagcgctt tggttggctc cttccgttgg ccaagcagtg caccactgct ttccatttcc 9060
tgttccggtg ttcaggagtt tggaggcaat aggtcaccgg tatgcagcaa tttccttggc 9120
tgtcataagt cactcgactc accaggcagc tgccaccttc agctgcaagc tactggcttg 9180
tgaagttatc atcttgtctt cggggcgtat gtggtttaag gagctgttta tgcgaagatg 9240
aaagcacttt tgtgcttcca agtcatttat tacatattgc gacatctatt acgttggaaa 9300
gtgtagtttt acaggggctg acaacattgt ctaaaacatg ttattaacgt atattgtaat 9360
attgtagtaa tgaagctcaa ttaaatcaag attattaatc tgtgcattta gaagggctgt 9420
taatttcgaa agatcctttt atgtcaagcc gtttttttctt ttatgcactg ctttattttg 9480
taatatctca acgcaattct atatggcgaa cgttctatcg taaggtagta aatgataatc 9540
tgagcacttt aagctccttt agcggtataa taaagacttt aatcaacatc taccccttgct 9600
gcagaaggca taaacgttgc gcaacattaa acggatgtcg tgggtgtacg tgtgggggcg 9660
tggccggggg cgctggcggt ggcagagcgc tcaatggtat gctacactta tttgcactgc 9720
actctggggc gtcctgtcat cgcgacgctg acatggcagc gatgggcttg ggatcggctt 9780
gtgtgccgca tttaacttcg ttttgttctt tattattcca cttgtcgctc atttaaatta 9840
catttatcgc gatggcgaaa aggattcgcc ctgccgcgtc actgtttgtt tgccatttcc 9900
attgtcctgc gacccttgcg gcaatgaccc tcaattaaac tcagttacag gtggtgaaaa 9960
caaaagtaat tggatataga cttaccgcag ctttcgattg cgctgcactg aaatgatgac 10020

caatgcattg ccgaggacag cggccaggat aattgacgag aagatgaatc ctttgagcag 10080
cagcagcgat agatccagcc agctctcctc cagctccgcc tggttatcca ccgcaatgtc 10140
actgccattg gccgttatat ttgggccgac tgtcgagagg agctcaaacg tggaatccga 10200
cgacaagtag gatgaaaagg gcgaggagta ggaggcatct gacgatgtgg ctgttaatga 10260
tagcgaagtc aacgaggcat ccacaaattg cgatgttata ggcaaactgc tcgtcggcgt 10320
cgtcgtagtt ttggctgctg ctgtcgtgga aatgtggcct gttaatgttg caggttgcag 10380
gtggctggcg ttggtggctg atgttgctgc tgcttgtgat gttgctgctc ctgctgcggc 10440
tgttatgggt aatgtcattg tcgtggccag aaggtcagcc acgttaaccc ctgacatctc 10500
gtcgatttcg cagcccaatt gagtcgcagc tcattcaagt tgctccgcca atgctttttc 10560
gctgtacttg tcgtttatct ttttcccggc tgctgtgcgt aatttaaaac tcgagtaatt 10620
gcaatctgca atgcaaacaa ggctgccgaa gtttgccgtc atcctggttg ccaatgttgc 10680
tcctctgcga ttccagtatt tgctcacagc tccgactgaa ggtctctgct ccatttccca 10740
atcctccata ggacgctttt cccgtttgag ttgctggaaa ttggcaaaaa gtcgagactc 10800
gcgcttcgat ttggcactca agcgcgcaac agtcggcgac acgggaactg caaatgaata 10860
aataaataat ttatgaaata atcggcgagt gtgaggctgt gaaactgcat ggcaaaattg 10920
gtaaaatggg aattcgaacg gttgcgccgc ccggggtgtt tttatgtaca tacgcaatat 10980
tcaacagggg gcataaaccc tattgcctat ttatgtgccg cagtaacaca taaatgggat 11040
tcttctgttt ctccagtgcc ccgtcaaatg tccgttcgtt aaaagtgcgc cacttcaact 11100
ttgctacact agattaaaat aacagaaata tattgcgtat acgcagtgtg tgcactcggc 11160
atgcatctgc atttctatat ccccagcgag acagcggtgc tgtcaagtgc tgacacttgt 11220
gcagagtgga atgtggaata tggacaagtt gttgctttaa aagtgagatt aaagcaagat 11280
tatcagtggg aaagttgttt cagtagtgcc tcacgcctgg gcgcagtgaa aactgcaaat 11340
ccttttttgtc ccggtacttg taagtaaaac aaaagactct tcataatctg tggtaatctc 11400
gaattttttc tttttttatta ttttaaactt tttgcgactc aatttggccg taaaaatgct 11460
aacacatcat cagggaaaag ctgaaagcgt gcttacactt tactttctcc cccgcaacat 11520
tgaattgaag ctgaatccgt ccttggcact cagacaaatc gctggctaca cgttttcatt 11580
gcacttattc tgtgcactct agatgccctg ccctcattcc aatcttcaga gttttgaatc 11640
taatattttt aaggatttat gtacttatac agataacctt gccgaattga ttgcattgcg 11700
ttgtccttga gctgtcgatt cgacttgttt actcagacca tccggtcaca tcctcctgct 11760
agattggcaa cttaatgttt taatgctcgg acttaaacca catcagtgcg ctctatcatc 11820
cttggcttat acatttcaca agccatt 11847

<210> 188
<211> 2092
<212> DNA
<213> Drosophila

<400> 188
ttcccgtgtc gccgactgtt gcgcgcttga gtgccaaatc gaagcgcgag tctcgacttt 60
ttgccaattt ccagcaactc aaacgggaaa agcgtcctat ggaggattgg gaaatggagc 120
agagaccttc agtcggagct gtgagcaaat actggaatcg cagaggagca acattggcaa 180
ccaggatgac ggcaaacttc ggcagccttg tttgcattgc agattgcaat tactcgagtt 240
ttaaattacg cacagcagcc gggaaaaaga taaacgacaa atgtcagggg ttaacgtggc 300
tgaccttctg gccacgacaa tgacattacc cataacagcc gcagcaggag cagcaacatc 360
acaagcagca gcaacatcag ccaccaacgc cagccacctg caacctgcaa cattaacagg 420
ccacatttcc acgacagcag cagccaaaac tacgacgacg ccgacgagca gtttgcctat 480
aacatcgcaa tttgtggatg cctcgttgac ttcgctatca ttaacagcca catcgtcaga 540
tgcctcctac tcctcgccct tttcatccta cttgtcgtcg gattccacgt ttgagctcct 600
ctcgacagtc ggcccaaata taacggccaa tggcagtgac attgcggtgg ataaccaggc 660
ggagctggag gagagctggc tggatctatc gctgctgctg ctcaaaggat tcatcttctc 720
gtcaattatc ctggccgctg tcctcggcaa tgcattggtc atcatttcag tgcagcgcaa 780
tcgaaagctg cgggtgataa ccaattactt tgttgtgtcg ctggcgatgg ccgacatgct 840
ggtggccctc tgtgcgatga cattcaacgc ctccgtggaa ctgtccggcg gaaagtggat 900
gttcggaccg ttcatgtgca acgtgtacaa cagcctggat gtttactttt ccacggccag 960
catattgcac ctgtgctgca tatcagtgga cagatactac gccattgtgc gtccactgga 1020
gtatccattg aatatgacac acaaaacggt ctgcttcatg ctcgccaatg tgtggatcct 1080
gcccgccctc atatccttca cgcccatctt tctgggctgg tacacgacgg aggagcacct 1140
gcgggagatt tccctgcatc cggaccagtg ctcgtttgtg gtcaacaagg cctacgccct 1200

```
catctccagt tcggtgagct tctggatacc cggcattgtg atgctggtga tgtactggcg 1260
catctttaag gaggcgattc gtcaacgcaa ggccctcagc cgcaccagct ccaacatcct 1320
ccttaacagc gtccacatgg gccacaccca gcagcccacc agcctgagct atctgcatcc 1380
cagtgactgt gatctgaatg ccacatccgc ccgggaggag acgcacagtg cgcttagtaa 1440
cttggagcgg ccccgcggca tcaccttcgc cgcggtgccc agtccgctgc ccaagtgtcc 1500
gctctgcggc gccgacatca gcagcaccac cggcaccacc gcaaacgcaa ccgcaaccgc 1560
caacgccgac agcacaatcg acaccacagt aaccactagt agtaaacgta gtattcacga 1620
gcaaactcca gatcttggcc agaggcccgc ctcctcgtcc agttccactc gattctggca 1680
caagcgaacg gcggcggtca cagcatgttg gcagcagagc aagaaccgga agcgccgctt 1740
caaaaccgga tgcagtcact gtggtgcaac cggcggttcc gtgcgaccgg ccaaaggatg 1800
gaaggccgaa cacaaggccg cccgcaccct gggcatcatc atgggcgtct ttctgctctg 1860
ctggctgccc ttctttctgt ggtatgtcat cacatcgctc tgcggtccgg cctgcccatg 1920
tcccgatgtg ctcgtggtgg tgttattctg gatcggttac ttcaactcca cgctaaatcc 1980
gcttatatac gcctacttta atcgcgattt tcgcgaggca ttccgcaata cgctggagtg 2040
tgtgctgccc tgtctggaga aacgaaatcc gtacaacgcc tactacgtct ag        2092
```

<210> 189
<211> 603
<212> PRT
<213> Drosophila

<400> 189

```
Met Ser Gly Val Asn Val Ala Asp Leu Leu Ala Thr Thr Met Thr Leu
 1               5                  10                  15
Pro Ile Thr Ala Ala Ala Gly Ala Ala Thr Ser Gln Ala Ala Ala Thr
            20                  25                  30
Ser Ala Thr Asn Ala Ser His Leu Gln Pro Ala Thr Leu Thr Gly His
        35                  40                  45
Ile Ser Thr Thr Ala Ala Ala Lys Thr Thr Thr Thr Pro Thr Ser Ser
    50                  55                  60
Leu Pro Ile Thr Ser Gln Phe Val Asp Ala Ser Leu Thr Ser Leu Ser
65                  70                  75                  80
Leu Thr Ala Thr Ser Ser Asp Ala Ser Tyr Ser Ser Pro Phe Ser Ser
                85                  90                  95
Tyr Leu Ser Ser Asp Ser Thr Phe Glu Leu Leu Ser Thr Val Gly Pro
            100                 105                 110
Asn Ile Thr Ala Asn Gly Ser Asp Ile Ala Val Asp Asn Gln Ala Glu
        115                 120                 125
Leu Glu Glu Ser Trp Leu Asp Leu Ser Leu Leu Leu Leu Lys Gly Phe
    130                 135                 140
Ile Phe Ser Ser Ile Ile Leu Ala Ala Val Leu Gly Asn Ala Leu Val
145                 150                 155                 160
Ile Ile Ser Val Gln Arg Asn Arg Lys Leu Arg Val Ile Thr Asn Tyr
                165                 170                 175
Phe Val Val Ser Leu Ala Met Ala Asp Met Leu Val Ala Leu Cys Ala
            180                 185                 190
Met Thr Phe Asn Ala Ser Val Glu Leu Ser Gly Gly Lys Trp Met Phe
        195                 200                 205
Gly Pro Phe Met Cys Asn Val Tyr Asn Ser Leu Asp Val Tyr Phe Ser
    210                 215                 220
Thr Ala Ser Ile Leu His Leu Cys Cys Ile Ser Val Asp Arg Tyr Tyr
225                 230                 235                 240
Ala Ile Val Arg Pro Leu Glu Tyr Pro Leu Asn Met Thr His Lys Thr
                245                 250                 255
Val Cys Phe Met Leu Ala Asn Val Trp Ile Leu Pro Ala Leu Ile Ser
            260                 265                 270
Phe Thr Pro Ile Phe Leu Gly Trp Tyr Thr Thr Glu Glu His Leu Arg
        275                 280                 285
Glu Ile Ser Leu His Pro Asp Gln Cys Ser Phe Val Val Asn Lys Ala
    290                 295                 300
```

```
Tyr Ala Leu Ile Ser Ser Ser Val Ser Phe Trp Ile Pro Gly Ile Val
305                 310                 315                 320
Met Leu Val Met Tyr Trp Arg Ile Phe Lys Glu Ala Ile Arg Gln Arg
                325                 330                 335
Lys Ala Leu Ser Arg Thr Ser Ser Asn Ile Leu Leu Asn Ser Val His
            340                 345                 350
Met Gly His Thr Gln Gln Pro Thr Ser Leu Ser Tyr Leu His Pro Ser
        355                 360                 365
Asp Cys Asp Leu Asn Ala Thr Ser Ala Arg Glu Glu Thr His Ser Ala
    370                 375                 380
Leu Ser Asn Leu Glu Arg Pro Arg Gly Ile Thr Phe Ala Ala Val Pro
385                 390                 395                 400
Ser Pro Leu Pro Lys Cys Pro Leu Cys Gly Ala Asp Ile Ser Ser Thr
                405                 410                 415
Thr Gly Thr Thr Ala Asn Ala Thr Ala Thr Ala Asn Ala Asp Ser Thr
            420                 425                 430
Ile Asp Thr Thr Val Thr Thr Ser Ser Lys Arg Ser Ile His Glu Gln
        435                 440                 445
Thr Pro Asp Leu Gly Gln Arg Pro Ala Ser Ser Ser Ser Ser Thr Arg
    450                 455                 460
Phe Trp His Lys Arg Thr Ala Ala Val Thr Ala Cys Trp Gln Gln Ser
465                 470                 475                 480
Lys Asn Arg Lys Arg Arg Phe Lys Thr Gly Cys Ser His Cys Gly Ala
                485                 490                 495
Thr Gly Gly Ser Val Arg Pro Ala Lys Gly Trp Lys Ala Glu His Lys
            500                 505                 510
Ala Ala Arg Thr Leu Gly Ile Ile Met Gly Val Phe Leu Leu Cys Trp
        515                 520                 525
Leu Pro Phe Phe Leu Trp Tyr Val Ile Thr Ser Leu Cys Gly Pro Ala
    530                 535                 540
Cys Pro Cys Pro Asp Val Leu Val Val Val Leu Phe Trp Ile Gly Tyr
545                 550                 555                 560
Phe Asn Ser Thr Leu Asn Pro Leu Ile Tyr Ala Tyr Phe Asn Arg Asp
                565                 570                 575
Phe Arg Glu Ala Phe Arg Asn Thr Leu Glu Cys Val Leu Pro Cys Leu
            580                 585                 590
Glu Lys Arg Asn Pro Tyr Asn Ala Tyr Tyr Val
        595                 600
```

<210> 190
<211> 6617
<212> DNA
<213> Drosophila

<400> 190

```
gttccgttac ttggtctttg acttcttacc ccagaatatg attcttggca gcaggagaaa 60
ccagtacgac ttgttttggt cgtattatga gtgaaatatt accactcatg tcattttgaa 120
cgtctttaaa gctcacaccc ttatggaccc attataagaa tatggcgtgg caaaatttcc 180
catttttatc gaggctggaa aaacttttcc acacatgcaa tcatgtttcg tttcacttat 240
catggctaaa acctagcaaa tagaaaaagt tttttttat ccacaaacat atctgtattg 300
agtgccagca atctgtattg ctttagctta agatgtttac tattaataga gccaagtggc 360
aaatgacata cggcttacag agaaaataat ttccaattca attacattta cgtacatata 420
ttccttccta caagcagaca ggattattga tctaaatata aataaatata taataattat 480
tttaataata taatattaaa atgtagatta acagaatata ttttttttgt atgaaagcct 540
ctagggtatt tgtgtgactt ttttctctct gcataagtga agatgtttac tattaataga 600
gccaagaggc aagagaaaag tggcagcaat caagtgcttt tatttcggga ctgcatgtta 660
aaggcgattt cctgatcgaa gactttggat acctgatgcc atccgtctga atgtgacgtc 720
agcctggctg gggattcaat ttaaattcat ccttgagtct aatgactttt aatgttgttt 780
gcttcacagg cgagtgcgca atttttgggc ctgtcgtcga ctgcttatcg gctatgtgcc 840
```

accgcagccg caaaggatac gacagacaat attaataacg cattgatacg agtgggcggc 900
cgatttgctt ttgtttttct ttcccaactt ggcggtggaa aggagtggtg tggcgtggcg 960
tggagtggat agagtccaca attaatcgac gacagctagt atgttgcaag gcgtcgccat 1020
caccatagcc aacgacagca acgatgatgg catcaatcaa tcattcatgg ctcatgtgtc 1080
gcccagtccc aatcagagtc cctccattgg agtgggcatc ggtatcgcct catccacaat 1140
ggcgaatccc agcgagagcc ccgagatgtt gctgctgaag aacgacaaat tcctaacaca 1200
tgtcgcccat ctgctgaaca taacgaccga gaatctttca aatctcctgg gttccacgaa 1260
cggcacaaat gcgagcacaa tggcagcgga ttcgccggtc gacgagtccc tgaccctgcg 1320
aacagccctc accgtttgtt atgcactcat ttttgtggcc ggcgtgctgg gtaatcttat 1380
cacctgcatt gttatatcgc gaaataactt tatgcacacg gccaccaatt tctatttgtt 1440
caacctggcg gtttccgact taatactgtt ggtctcaggt gagatggcat tttgagggga 1500
caccattgcg tatgcgtaat gcggcaaata tcctgcaggc atcccccagg agctgtacaa 1560
cctttggtat ccggatatgt atcccttcac ggacgccatg tgcatcatgg gcagtgtgct 1620
ctcggaaatg gccgccaatg caacggttct caccatcaca gcattcactg tggagcgata 1680
tatagcaatc tgtcatccat tccggtgagt ttttaacatt taaacgttat atatatttta 1740
aatatttatt tacattatta aacaaatgta taatttcttc agtgttattg cctcgctctc 1800
attaacaaag tggatcatgc ataattagga cagagaatat gaatagataa attaaacttg 1860
ataagcgtca gtcggaaaca ttatgatatg aaacaacaaa agtacctcac aaattatggg 1920
cacaataatt attcataact aaacagtctg gctgtggaat gataaaacaa gttaggcttg 1980
acaaccaaat tcccatcgtt tttaaatata ccaaattcat tagctttcat ttatagacac 2040
taattaaagt caggtccata aatgaaccaa gttcattaca gactatgtaa ttaagttcca 2100
ataaatgcgt tgtaaagata gaaagttagt tagtcaatta attgcttagt accaacatat 2160
ttgcattttt ggaatggtac aacaaaactt ttccttatac aggtatcacc agttccttcc 2220
aaatgttctg accattttgc ggtcaaagaa tacaacagct tccttaagcg aaagcgtctg 2280
gttttcattt acgatgttgc atttccttta gttcagaaac tcttgaatag ctagaagaaa 2340
tgaaacttaa aaaccttaca ttgttgtttt caaacacaat ttccgtttct tagcattttt 2400
cacaaattaa ttgttacgaa tacaatcact tggagtatgt ccttaagaat atgcagacac 2460
aaaagggtgg ttttctataa gaagggattc caagtccgat tacaaagtta aggcaaaaaa 2520
ctttcatgtc tctgaagttg ggattaaata ttgttgtcga aaagagtaaa aggtttctat 2580
gtttgccatt taatagtatt atagttattt tgaaacgaaa ggatattcgg acatgtgctt 2640
agcatgaaaa aggtctgaaa cccattgcag tcactttacc aaaagaagtt aaaagtgatg 2700
tctattaaat ttggtatgaa actgtttgaa ttccttccca tttacaccca cattcaaagt 2760
aaatatttat gttgcactca atatcccaaa ccttaaagag aaagttaaaa agcgtaaaaa 2820
gtgtaggtgg tcactgcgtg tactttacaa ccgtaatatt ttccatctta tatctgccca 2880
cgtttatccg aaacatcgac atcatcttgt ccttaatact taacatataa tgtagcctaa 2940
tctttccgtc atttgtaggc agcacaccat gtcaaaattg tcgcgagcca ttaaatttat 3000
atttgccatt tggctggcgg ccttcctttt ggcccttccg caggccatgc aattttcggt 3060
ggtctaccag aacgaaggat actcgtgcac ggtgggtggg actgaaatgt gaaaaccaac 3120
tgacgcatca tcgagtgtcc ttttcgctgt ttattccctt agatggaaaa cgacttttac 3180
gcccatgtgt tcgccgtttc gggatttatt ttcttcggtg gacccatgac ggcgatttgt 3240
gtactgtacg tgctgatcgg agtgaaactg aagaggagtc gactgctgca atcgctgccg 3300
cgaaggactt tcgatgcgaa tcgcggccta aatgcccagg gacgagtcat cagaatgttg 3360
ggtaggtgag ttaagttggc ggcttgattg attaaagaca gcaatatgcc ccaccgccaa 3420
ccagaaacaa caagaacatc cacggctgtc aaaacttttt ccgccgaaag tcttaaattg 3480
aaaattactg tcaacgtatt tgcgttggcg aacatcaaac agggaattcg tcatagcccc 3540
gccctcattg ccttccgccc acttcatttg cacttgcgtc atgtgggctg aacgtgcttc 3600
tgttttttc cctcttttcc ctcaacggtg tcgttaattt ttcgcaaaca tcggtgcgta 3660
tgtgtaaagc tcttaaatgg ccatctagcc tatgaagcaa aggtgaagcc aggtgtcacc 3720
tcaaaagtaa tcaagccggg gcacttaaat gaactcaaga ctgctgggtg cacacagcaa 3780
gcttaaatat ttgtaaaaag tgtttcttga atctatagac atatgtccgt taacgtgcca 3840
cgttataagc tgttaaattt taaaggaatt cagcacacat ctcgtgtatt attcacatat 3900
tttatttata taattagaaa ttcttatcaa accatacagt gaataaattt aagtgtaaat 3960
ccaggcccaa ttttagatcc caataagccg ttttttctta cagcctacga aacaacgttt 4020
aatttactgt tgcttttaac tgaaacaccg accattgatt tcgaaattta ccattggccc 4080
tggcctgcca aatggaaatt gatggaccaa ttcagattaa gcccagagcc ttaaccataa 4140
attaatttgg ccagcaagag tggaaaaaaa acacacccgt tctcagggct tagttattca 4200
tagatagtgg gatattattt atcgtttgct aagaggtaat accttggctt ttatcaaaat 4260
gtgatatttt accaaaggcc cgtttgcgca aactattcga tttcgggggа tataaaagat 4320
taagaccact ttcagtgtga gagagcccct tttccggtga cctagttgtt caggggtctt 4380
aatacgttgc ttatccttac ttgaccgtat ttgcagtagc tgtagccgtc gccttcttcc 4440

tctgctgggc tccattccac gcccagcgac tgatggcagt ttacggactg aatctgatta 4500
atataggcat cagtcgggac gccttcaacg attacttccg gatacttgat tacacgtccg 4560
gggtgctcta ctttctctcc acctgcatca atccgctact gtacaacatc atgagccaca 4620
agttccgcga ggccttcaag gtaagcccat gggatcaacc cttctcagtt aaccacttca 4680
taatagtaat ctattattaa ttgggtcctt cgcatgtttg tgaggacggt cacaaatgca 4740
gatgctttta aagttaaatt taagctttaa ttattcaaaa ttgtttttcg taaattaatg 4800
aattggtaat ctatgtattt tttacacttt taagatcact ctgacgcgac agtttggtct 4860
ggccaggaac caccaccacc agcagagtca gcaccaccag cacaactaca gtgcccttct 4920
ccgccagaat ggatcaatgc gtctgcagcc ggccagttgc agtgtaaaca acaacgcgct 4980
ggagccatat ggatcctacc gagtggttca gtttcgttgc cgggatgcga accaccaatt 5040
gtccctgcag gacagcattc gcaccaccac cacaaccacg acgataaaca gcaacagcat 5100
ggcagccgga aatggagttg gtggtggtgc tggtggcggt ggcggtgggc ggcgtctgcg 5160
gaaacaggag ctctatgggc caggtccggg tacagcagtt ccccatcgaa tgctgcaggc 5220
ccaggtgtct caactctcgt cgctgggcga tgccaattca ctgctggagg cggaggtggt 5280
ggacagacac tacgcatccg ggcgagccaa aagagctctc ttggccacca aaagtggtgc 5340
cctgctggtg acaccacctc aaagtggtga cccatcggaa gtcagtcagc ccgccactcg 5400
tctcaagctg actagggtga taagtagaag ggatgaggtg gccaacacaa gcactccgcc 5460
cttttgtggg agccacagtt tgccggatcc ggagacttgt caatcggcat cggtggctgg 5520
tcgaagctcc cgtaagtttc cgtggcgaaa acgaaggcaa aaaacggagg atcccagcag 5580
cgagggcttg acctatggct cacccaaatc ccagtgattt cagctggctg gtagctagag 5640
ttatacttag ttaggcattg tttagttagc gcaataaact tttctatctg taaataagcc 5700
agaaaaagtc ctgcgatttg aatgcaaatt tcgggtgacc ataaaaaatt taacgaatgg 5760
gcatggtcca gacaatttat gagagtcata aagccgggat aatgaggcta cggcgtatcg 5820
tgacgaatac acgcccccaca tcattttgca atggcaaccg catagcactc ataaattaag 5880
tgttaagcta ccagctaaat aatgtgaata aataaaaatt gttctcaata ttacagcaaa 5940
atccttttaa atgcgttttc ctttcacaat gcatgccagt tacattcaag ttcgcctgca 6000
ttccaagttg tccaggatca aagggaaaat caaaaaaaac agtttacatc tcaagttttg 6060
caggtgcagc gtttccaagc aatctgttcc taccttttgc gacatttttc ctgcagattg 6120
cctggcattc tgttgatttc ccacacaaaa actaaaagcg ctgaaaaaaa taaaagaagc 6180
ataaaataaa ataaagtgcc aataaattag cagtcttaat attttcagag cccctcagaa 6240
gcaaatctag gaaggaaaac atgacataaa ttttaaatgg attgctcatc aggatatcgt 6300
gttgcagtag atatatatat tttcgcagtg taagtcacga ctccacttgc acctggcaac 6360
aaagtttcag tttgcacctg gcccattgcc ggaggtccat tagtcaagag ggctttgtag 6420
gctaggaagt ccggattgaa acgaatgtgc agagagtggc acgaatgggc ttatcctgcg 6480
gtgaattggc acgtggagtg catggagagt ggtgtcggtg ccacaaaata aatcaactta 6540
aaagtagcat acaaatttaa ggacccgact gatgttatgt ccaaagggaa ttattaatgg 6600
ttctttaaag gcaggca 6617

<210> 191
<211> 1983
<212> DNA
<213> Drosophila

<400> 191
atgttgcaag gcgtcgccat caccatagcc aacgacagca acgatgatgg catcaatcaa 60
tcattcatgg ctcatgtgtc gcccagtccc aatcagagtc cctccattgg agtgggcatc 120
ggtatcgcct catccacaat ggcgaatccc agcgagagcc ccgagatgtt gctgctgaag 180
aacgacaaat tcctaacaca tgtcgcccat ctgctgaaca taacgaccga gaatctttca 240
aatctcctgg gttccacgaa cggcacaaat gcgagcacaa tggcagcgga ttcgccggtc 300
gacgagtccc tgaccctgcg aacagccctc accgtttgtt atgcactcat ttttgtggcc 360
ggcgtgctgg gtaatcttat cacctgcatt gttatatcgc gaaataactt tatgcacacg 420
gccaccaatt tctatttgtt caacctggcg gtttccgact taatactgtt ggtctcaggc 480
atcccccagg agctgtacaa cctttggtat ccggatatgt atcccttcac ggacgccatg 540
tgcatcatgg gcagtgtgct ctcggaaatg gccgccaatg caacggttct caccatcaca 600
gcattcactg tggagcgata tatagcaatc tgtcatccat tccggcagca caccatgtca 660
aaattgtcgc gagccattaa atttatattt gccatttggc tggcggcctt ccttttggcc 720
cttccgcagg ccatgcaatt ttcggtggtc taccagaacg aaggatactc gtgcacgatg 780
gaaaacgact tttacgccca tgtgttcgcc gtttcgggat ttattttctt cggtggaccc 840

```
atgacggcga tttgtgtact gtacgtgctg atcggagtga aactgaagag gagtcgactg 900
ctgcaatcgc tgccgcgaag gactttcgat gcgaatcgcg gcctaaatgc ccagggacga 960
gtcatcagaa tgttggtagc tgtagccgtc gccttcttcc tctgctgggc tccattccac 1020
gcccagcgac tgatggcagt ttacggactg aatctgatta atataggcat cagtcgggac 1080
gccttcaacg attacttccg gatacttgat tacacgtccg ggtgctcta ctttctctcc 1140
acctgcatca atccgctact gtacaacatc atgagccaca agttccgcga ggccttcaag 1200
atcactctga cgcgacagtt tggtctggcc aggaaccacc accaccagca gagtcagcac 1260
caccagcaca actacagtgc ccttctccgc cagaatggat caatgcgtct gcagccggcc 1320
agttgcagtg taaacaacaa cgcgctggag ccatatggat cctaccgagt ggttcagttt 1380
cgttgccggg atgcgaacca ccaattgtcc ctgcaggaca gcattcgcac caccaccaca 1440
accacgacga taaacagcaa cagcatggca gccggaaatg gagttggtgg tggtgctggt 1500
ggcggtggcg gtgggcggcg tctgcggaaa caggagctct atgggccagg tccgggtaca 1560
gcagttcccc atcgaatgct gcaggcccag gtgtctcaac tctcgtcgct gggcgatgcc 1620
aattcactgc tggaggcgga ggtggtggac agacactacg catccgggcg agccaaaaga 1680
gctctcttgg ccaccaaaag tggtgccctg ctggtgacac cacctcaaag tggtgaccca 1740
tcggaagtca gtcagcccgc cactcgtctc aagctgacta gggtgataag tagaagggat 1800
gaggtggcca acacaagcac tccgcccttt tgtgggagcc acagtttgcc ggatccggag 1860
acttgtcaat cggcatcggt ggctggtcga agctcccgta agtttccgtg gcgaaaacga 1920
aggcaaaaaa cggaggatcc cagcagcgag ggcttgacct atggctcacc caaatcccag 1980
tga                                                               1983

&lt;210&gt; 192
&lt;211&gt; 660
&lt;212&gt; PRT
&lt;213&gt; Drosophila

&lt;400&gt; 192
Met Leu Gln Gly Val Ala Ile Thr Ile Ala Asn Asp Ser Asn Asp Asp
 1               5                  10                  15
Gly Ile Asn Gln Ser Phe Met Ala His Val Ser Pro Ser Pro Asn Gln
            20                  25                  30
Ser Pro Ser Ile Gly Val Gly Ile Gly Ile Ala Ser Ser Thr Met Ala
        35                  40                  45
Asn Pro Ser Glu Ser Pro Glu Met Leu Leu Leu Lys Asn Asp Lys Phe
    50                  55                  60
Leu Thr His Val Ala His Leu Leu Asn Ile Thr Thr Glu Asn Leu Ser
65                  70                  75                  80
Asn Leu Leu Gly Ser Thr Asn Gly Thr Asn Ala Ser Thr Met Ala Ala
                85                  90                  95
Asp Ser Pro Val Asp Glu Ser Leu Thr Leu Arg Thr Ala Leu Thr Val
            100                 105                 110
Cys Tyr Ala Leu Ile Phe Val Ala Gly Val Leu Gly Asn Leu Ile Thr
        115                 120                 125
Cys Ile Val Ile Ser Arg Asn Asn Phe Met His Thr Ala Thr Asn Phe
    130                 135                 140
Tyr Leu Phe Asn Leu Ala Val Ser Asp Leu Ile Leu Leu Val Ser Gly
145                 150                 155                 160
Ile Pro Gln Glu Leu Tyr Asn Leu Trp Tyr Pro Asp Met Tyr Pro Phe
                165                 170                 175
Thr Asp Ala Met Cys Ile Met Gly Ser Val Leu Ser Glu Met Ala Ala
            180                 185                 190
Asn Ala Thr Val Leu Thr Ile Thr Ala Phe Thr Val Glu Arg Tyr Ile
        195                 200                 205
Ala Ile Cys His Pro Phe Arg Gln His Thr Met Ser Lys Leu Ser Arg
    210                 215                 220
Ala Ile Lys Phe Ile Phe Ala Ile Trp Leu Ala Ala Phe Leu Leu Ala
225                 230                 235                 240
Leu Pro Gln Ala Met Gln Phe Ser Val Val Tyr Gln Asn Glu Gly Tyr
                245                 250                 255
Ser Cys Thr Met Glu Asn Asp Phe Tyr Ala His Val Phe Ala Val Ser
```

```
            260                     265                     270
Gly Phe Ile Phe Phe Gly Gly Pro Met Thr Ala Ile Cys Val Leu Tyr
        275                     280                     285
Val Leu Ile Gly Val Lys Leu Lys Arg Ser Arg Leu Leu Gln Ser Leu
    290                     295                     300
Pro Arg Arg Thr Phe Asp Ala Asn Arg Gly Leu Asn Ala Gln Gly Arg
305                     310                     315                     320
Val Ile Arg Met Leu Val Ala Val Ala Val Ala Phe Phe Leu Cys Trp
                325                     330                     335
Ala Pro Phe His Ala Gln Arg Leu Met Ala Val Tyr Gly Leu Asn Leu
            340                     345                     350
Ile Asn Ile Gly Ile Ser Arg Asp Ala Phe Asn Asp Tyr Phe Arg Ile
        355                     360                     365
Leu Asp Tyr Thr Ser Gly Val Leu Tyr Phe Leu Ser Thr Cys Ile Asn
    370                     375                     380
Pro Leu Leu Tyr Asn Ile Met Ser His Lys Phe Arg Glu Ala Phe Lys
385                     390                     395                     400
Ile Thr Leu Thr Arg Gln Phe Gly Leu Ala Arg Asn His His His Gln
                405                     410                     415
Gln Ser Gln His His Gln His Asn Tyr Ser Ala Leu Leu Arg Gln Asn
            420                     425                     430
Gly Ser Met Arg Leu Gln Pro Ala Ser Cys Ser Val Asn Asn Asn Ala
        435                     440                     445
Leu Glu Pro Tyr Gly Ser Tyr Arg Val Val Gln Phe Arg Cys Arg Asp
    450                     455                     460
Ala Asn His Gln Leu Ser Leu Gln Asp Ser Ile Arg Thr Thr Thr Thr
465                     470                     475                     480
Thr Thr Thr Ile Asn Ser Asn Ser Met Ala Ala Gly Asn Gly Val Gly
                485                     490                     495
Gly Gly Ala Gly Gly Gly Gly Gly Gly Arg Arg Leu Arg Lys Gln Glu
            500                     505                     510
Leu Tyr Gly Pro Gly Pro Gly Thr Ala Val Pro His Arg Met Leu Gln
        515                     520                     525
Ala Gln Val Ser Gln Leu Ser Ser Leu Gly Asp Ala Asn Ser Leu Leu
    530                     535                     540
Glu Ala Glu Val Val Asp Arg His Tyr Ala Ser Gly Arg Ala Lys Arg
545                     550                     555                     560
Ala Leu Leu Ala Thr Lys Ser Gly Ala Leu Leu Val Thr Pro Pro Gln
                565                     570                     575
Ser Gly Asp Pro Ser Glu Val Ser Gln Pro Ala Thr Arg Leu Lys Leu
            580                     585                     590
Thr Arg Val Ile Ser Arg Arg Asp Glu Val Ala Asn Thr Ser Thr Pro
        595                     600                     605
Pro Phe Cys Gly Ser His Ser Leu Pro Asp Pro Glu Thr Cys Gln Ser
    610                     615                     620
Ala Ser Val Ala Gly Arg Ser Ser Arg Lys Phe Pro Trp Arg Lys Arg
625                     630                     635                     640
Arg Gln Lys Thr Glu Asp Pro Ser Ser Glu Gly Leu Thr Tyr Gly Ser
                645                     650                     655
Pro Lys Ser Gln
            660
```

<210> 193
<211> 4902
<212> DNA
<213> Drosophila

<400> 193
ttgggcaaat taaggaaact ttttgtggtt tcagttactc gtttaagcca cattttccca 60

acatatgctg cactttatga attaaaagtc atttatttta ggctgcacac acgttatgtt 120
tgcctgacaa attgcatttg ctgaaattaa aacaaaatgc actgggacac tgcgttatgg 180
catgctgcta attagtcacc aacagcggct ttcagcgaat gttggccaat cgcgagaatg 240
gctttcatcc actgtctaga ctccagacca aaaagcgaat ttgtgcgcca gataaacatc 300
ttaatggcca gggattatga tgcagtgccc gctttttact gacagtaatt agccagacgc 360
ccattgaacc ctcttcttga taatccccac ttgatccaca ctctttcaaa aaaaaaaaat 420
gtattcatta aatgggaaaa aaatatttac tgcaagagaa atgttattaa tggatattga 480
aaaactgaca ttagataaca gtttctattg gactgctaac aaagtcggtc agctgtttaa 540
aaaactgaga tcgcttttat atttaaattg tataataaaa ccattagggt atcaacctta 600
atatgtattt atatatcact gtactccaaa tactatttac gcttaactat gtagattcat 660
tttattttta tcaaaagtcc attctttttt tattaaactc cactttataa atgaaacaac 720
ggtcgaattt ttgaactgag tacataacta catagtttat tcgattacgt gcactataac 780
tataattcga actatagctt tgggctatat tacggttatt ggtaaggaca accatataat 840
tgaccccaaa cgcggaatat ttgtctaaat taaaacacaa acaagctgga gtggcgcttg 900
ctgctcgtat cacagttgtg tgtgtgtgta atgtacgtag gtggcactgg aatggactac 960
ctcgcttatc gagtacagat agaggatatg ctgacccccct ttaaaaggcg gcccgctctt 1020
cagtgtcgcc tccattcacc gtcgcccccct tttcgtgcgg tttgtggctg cctttgcgtc 1080
tggcacgcca attgaatttg gtcctgagac ccttgaacgg gcttgagctc ttgagcactt 1140
tcactttcgc cttgcgcact ttgccactcg actcgtcacc gctcacgccc cccgaccccg 1200
cggcccctgt tgtgctgcct cctgccacgc cccccgtcac acctcctggt gggcgggtga 1260
tgcgcgtaag tttaaggcgt gaccgcgaga tggccaagtc atctgtgagc tcgtcaatgg 1320
agcacatcgt gggctggcgc tgcagtggct ccccgggaaa ctgcaaatcc tcctccaaga 1380
gggaatgggc atcgcccacc gatgacagct gcgatatttg cgtctggagc atgcgacgcg 1440
gattatcgaa gaggtcttgc cggctgcgat tctgaccctg catctggaca ctgtccaagg 1500
acacgcggtt caggcgcacc gactgacccg ttgtcgttcc attgccactg ccattggcag 1560
agccatttag cccagtggtg gtcgtcgcca tggaagtcat tgtggtgcga acgctgtcct 1620
gcaggaggaa atcaaggaat ggagggataa tcagaaaggg aaagtgcagt cgccagacac 1680
gaaaataagt aacgcaataa acatcttata gtggttttaa aaagtctcgt gattcaaatt 1740
tcattttgag aaacaggaca tacactatat taatcctttt cattttcctc ttaaataatg 1800
ccacataata gaatttaagt gcatatatcg aatgataatt atgtaatttt tctcagtgca 1860
ctcaccgttg tgtgcagccg cagtgaaccc gtttgattac gtcgaagagc gctgtaagta 1920
tgaggaagcc cccttccctg atttttgcca ccaagtccaa agtgtctagc caaagttacc 1980
taaaatggga ttaaatacgt tttttaattc tagcttgaag tttaaaatca ctaaattgac 2040
cgttaaagaa tcgattaata aacttgtaac ctgaagctta ttttcaattt cctaacgcaa 2100
gctatagaaa tgaatatccg gcagcgaaac tcaccttaaa agcctcacga aacttgtggc 2160
tcatgatgtt gtagagcagc gggttgatgc aagttgagag gaagtagaga acaccggacg 2220
tatagtcgag gatgctgaac acgtcgttga accactggga ctcaatgccc gaggtggatc 2280
catagaccgc catcagccgc tgggcgtgaa agggggccca gcagatgaag aaggccacgg 2340
ccaccgccac tgggaaagtt ggataagggc gcctgatgca accatgagtt atttcggata 2400
tttcatatcc tgcttaccca gcatccggat gactcgcgtt tgggcgctta tcccccggtt 2460
tacatcgtaa catcgcctgg gaagcgcctg caggagtcgg ctccgtttca acttcacccc 2520
gatgaggaca tagagcacgc agatggccgt catgggtccg ccaaagaaca ggaagcccga 2580
cacagcaaac acatgggcaa aaaagtcgtt tttcatctac agcaggcatt tcacaaaaat 2640
aatatatatg tgtattcgca tgcgatacga tatggttttg gcttacagcc acagccggaa 2700
ggcaatggct gtaaatttga ttgcctaaac tgagtcgaac ctcgtaaggg attgtcgctg 2760
tttttatcat gctcttaaat tttttgtaat tcacaaaggc tctctattcg aactttctat 2820
acggaaaaca cataatgaat caaaggatca aaggaccggt tctacgaatc aattaaccag 2880
gtaactcata actgggggac atacgattaa gggtctcgca tcaatcatac gcaccgtgca 2940
cgatgttccc atgccctgca tcaccacgga gaactgaatg gcttggggca gggccagcaa 3000
aagggcagct atccagatgg caaatatgaa ctttacggcc cgtgacaact tggacatcgt 3060
gtgctgcctg tatcgaaagc agaaccaatt aagcaagttt ggggcaaaat ggaaatgggc 3120
agtggaagga atgaaggttc ccttacctga acggatgaca aatggcaata tatcgttcga 3180
ctgtgaacgc ggtaatggtt agaactgtcg cattggccgc cgtttccgag agaacgctct 3240
ccaatatgca gatgctgtca ctgaaaggat aattatccgg gtgccagagg ttatagaggt 3300
cctgcggcat tcctaagaac ataaatagca aaatagcaaa atatatagtg tgattataaa 3360
tatttaataa tagctgaaat tatagcctat agcctcaaaa tgcaaatcga tttgagcatc 3420
tgccgaatat ctcactctac attttcttgt ttatttacaa attttaattt aatatgcctc 3480
agttccattt attgtttgca ctactgaact gaattaaatt ataattcaaa aaccatgaac 3540
attttagttt tgcactcacc tgagcataac aaaatcatgt cggatatagc gaggttaaac 3600
agataaaagt tggtggccgt gtgcataaag ttgttccgcg aaatgacaat gcatgtgatg 3660

aggttgccca aaacgccggc gataaatatg agggcgtagc ccacgcttag ggtggccagc 3720
agtgaaagtg gtgtcatcgg ggtcaccgtg ggcaaaagct cctccggctc caggccctgg 3780
agaaggctgg tcaggttgtc cgcggagatg ttcagcactt gtgtcagatt taagagaaat 3840
ttctcattgt gaaacagttg tagatcggtt gccaggacac cggaactgtt agtgggcagc 3900
attttgactg ccatagggtg tgtgtctttg tgtgtgtgta tacgcactat tattttcgaa 3960
tttctttgga aatttcagac ttctgcgtat gtattctcaa tcgaagcttt cgcttgcaaa 4020
taatgttttc ataacgcccc ttacgggggc tcatcaattt cgacttaggc cagagccact 4080
ccacttgcaa tctgttgtca tgagtgtggg attaaattat gtgtacggat gtgtgtgata 4140
aggtggaccc tatgtgctat ttaaataaaa ctatgttttt caaaatttgt taaaatttaa 4200
ataaaataat acaagacaaa taacgcattt taacgaattt agtaataaaa caaattcaaa 4260
cattaaggct ttttactcca gatttctaaa ttaccataaa ttatacatat acacaatttg 4320
gttggtaagc tgtgcgcttt gtttagcttg ctaaggtcca ccttagagag cccgtcttaa 4380
agcgtgtttg tgtcctgtgt gtttacctgg catcctggga aactggcagg cggcagtcag 4440
gcgtgtccta ggcgaaagac caggatcagg gcgagtggaa gctgaaaatt aatttgaatt 4500
gctaagtggc atttagtcaa aatatttgtg tacactggat gaatttaagg atgttgctac 4560
aaggctgaga aggacaggcc tgccgatccg tctcccaccc aacaccaggt ggcaaacgaa 4620
tgaaacgagc cagaagctca aggtcaggtg gtggataaga cggataagga gacgacgatg 4680
gcaaatggac accgggtgac ttaataacat taacatagtt tcattccatt cagatgatga 4740
gtgttttcgg ctgctgctgt tctgctctgt tctgttcaca ttctggctat acttttcgtt 4800
attgttgctg ctgaccgctg ttgaccgttt ttggacgttt tcattatcat tatcataaaa 4860
aagtttctac aatatatgca caaccaaaag aatgctgtcg aa 4902

<210> 194
<211> 1788
<212> DNA
<213> Drosophila

<400> 194
atgctgccca ctaacagttc cggtgtcctg gcaaccgatc tacaactgtt tcacaatgag 60
aaatttctct taaatctgac acaagtgctg aacatctccg cggacaacct gaccagcctt 120
ctccagggcc tggagccgga ggagcttttg cccacggtga ccccgatgac accactttca 180
ctgctggcca ccctaagcgt gggctacgcc ctcatattta tcgccggcgt tttgggcaac 240
ctcatcacat gcattgtcat ttcgcggaac aactttatgc acacggccac caacttttat 300
ctgtttaacc tcgctatatc cgacatgatt ttgttatgct caggaatgcc gcaggacctc 360
tataacctct ggcacccgga taattatcct ttcagtgaca gcatctgcat attggagagc 420
gttctctcgg aaacggcggc caatgcgaca gttctaacca ttaccgcgtt cacagtcgaa 480
cgatatattg ccatttgtca tccgttcagg cagcacacga tgtccaagtt gtcacgggcc 540
gtaaagttca tatttgccat ctggatagct gccctttgc tggccctgcc ccaagccatt 600
cagttctccg tggtgatgca gggcatggga acatcgtgca cgatgaaaaa cgactttttt 660
gcccatgtgt ttgctgtgtc gggcttcctg ttctttggcg gacccatgac ggccatctgc 720
gtgctctatg tcctcatcgg ggtgaagttg aaacggagcc gactcctgca ggcgcttccc 780
aggcgatgtt acgatgtaaa ccgggggata agcgcccaaa cgcgagtcat ccggatgctg 840
gtggcggtgg ccgtggcctt cttcatctgc tgggccccct ttcacgccca gcggctgatg 900
gcggtctatg gatccacctc gggcattgag tcccagtggt tcaacgacgt gttcagcatc 960
ctcgactata cgtccggtgt tctctacttc ctctcaactt gcatcaaccc gctgctctac 1020
aacatcatga gccacaagtt tcgtgaggct tttaaggtaa ctttggctag acactttgga 1080
cttggtggca aaaatcaggg aaggggcttt cctcatactt acagcgctct tcgacgtaat 1140
caaacgggtt cactgcggct gcacacaacg gacagcgttc gcaccacaat gacttccatg 1200
gcgacgacca ccactgggct aaatggctct gccaatggca gtggcaatgg aacgacaacg 1260
ggtcagtcgg tgcgcctgaa ccgcgtgtcc ttggacagtg tccagatgca gggtcagaat 1320
cgcagccggc aagacctctt cgataatccg cgtcgcatgc tccagacgca aatatcgcag 1380
ctgtcatcgg tgggcgatgc ccattccctc ttggaggagg atttgcagtt tcccggggag 1440
ccactgcagc gccagcccac gatgtgctcc attgacgagc tcacagatga cttggccatc 1500
tcgcggtcac gccttaaact tacgcgcatc acccgcccac caggaggtgt gacgggggc 1560
gtggcaggag gcagcacaac aggggccgcg gggtcggggg gcgtgagcgg tgacgagtcg 1620
agtggcaaag tgcgcaaggc gaaagtgaaa gtgctcaaga gctcaagccc gttcaagggt 1680
ctcaggacca aattcaattg gcgtgccaga cgcaaaggca gccacaaacc gcacgaaaag 1740
gggcgacgg tgaatggagg cgacactgaa gagcgggccg ccttttaa 1788

<210> 195

```
<211> 595
<212> PRT
<213> Drosophila

<400> 195
Met Leu Pro Thr Asn Ser Ser Gly Val Leu Ala Thr Asp Leu Gln Leu
 1               5                  10                  15
Phe His Asn Glu Lys Phe Leu Leu Asn Leu Thr Gln Val Leu Asn Ile
            20                  25                  30
Ser Ala Asp Asn Leu Thr Ser Leu Leu Gln Gly Leu Glu Pro Glu Glu
        35                  40                  45
Leu Leu Pro Thr Val Thr Pro Met Thr Pro Leu Ser Leu Leu Ala Thr
    50                  55                  60
Leu Ser Val Gly Tyr Ala Leu Ile Phe Ile Ala Gly Val Leu Gly Asn
65                  70                  75                  80
Leu Ile Thr Cys Ile Val Ile Ser Arg Asn Asn Phe Met His Thr Ala
                85                  90                  95
Thr Asn Phe Tyr Leu Phe Asn Leu Ala Ile Ser Asp Met Ile Leu Leu
            100                 105                 110
Cys Ser Gly Met Pro Gln Asp Leu Tyr Asn Leu Trp His Pro Asp Asn
        115                 120                 125
Tyr Pro Phe Ser Asp Ser Ile Cys Ile Leu Glu Ser Val Leu Ser Glu
    130                 135                 140
Thr Ala Ala Asn Ala Thr Val Leu Thr Ile Thr Ala Phe Thr Val Glu
145                 150                 155                 160
Arg Tyr Ile Ala Ile Cys His Pro Phe Arg Gln His Thr Met Ser Lys
                165                 170                 175
Leu Ser Arg Ala Val Lys Phe Ile Phe Ala Ile Trp Ile Ala Ala Leu
            180                 185                 190
Leu Leu Ala Leu Pro Gln Ala Ile Gln Phe Ser Val Val Met Gln Gly
        195                 200                 205
Met Gly Thr Ser Cys Thr Met Lys Asn Asp Phe Phe Ala His Val Phe
    210                 215                 220
Ala Val Ser Gly Phe Leu Phe Phe Gly Gly Pro Met Thr Ala Ile Cys
225                 230                 235                 240
Val Leu Tyr Val Leu Ile Gly Val Lys Leu Lys Arg Ser Arg Leu Leu
                245                 250                 255
Gln Ala Leu Pro Arg Arg Cys Tyr Asp Val Asn Arg Gly Ile Ser Ala
            260                 265                 270
Gln Thr Arg Val Ile Arg Met Leu Val Ala Val Ala Val Ala Phe Phe
        275                 280                 285
Ile Cys Trp Ala Pro Phe His Ala Gln Arg Leu Met Ala Val Tyr Gly
    290                 295                 300
Ser Thr Ser Gly Ile Glu Ser Gln Trp Phe Asn Asp Val Phe Ser Ile
305                 310                 315                 320
Leu Asp Tyr Thr Ser Gly Val Leu Tyr Phe Leu Ser Thr Cys Ile Asn
                325                 330                 335
Pro Leu Leu Tyr Asn Ile Met Ser His Lys Phe Arg Glu Ala Phe Lys
            340                 345                 350
Val Thr Leu Ala Arg His Phe Gly Leu Gly Gly Lys Asn Gln Gly Arg
        355                 360                 365
Gly Leu Pro His Thr Tyr Ser Ala Leu Arg Arg Asn Gln Thr Gly Ser
    370                 375                 380
Leu Arg Leu His Thr Thr Asp Ser Val Arg Thr Thr Met Thr Ser Met
385                 390                 395                 400
Ala Thr Thr Thr Thr Gly Leu Asn Gly Ser Ala Asn Gly Ser Gly Asn
                405                 410                 415
Gly Thr Thr Thr Gly Gln Ser Val Arg Leu Asn Arg Val Ser Leu Asp
            420                 425                 430
Ser Val Gln Met Gln Gly Gln Asn Arg Ser Arg Gln Asp Leu Phe Asp
```

```
        435                 440                 445
Asn Pro Arg Arg Met Leu Gln Thr Gln Ile Ser Gln Leu Ser Ser Val
    450                 455                 460
Gly Asp Ala His Ser Leu Leu Glu Glu Asp Leu Gln Phe Pro Gly Glu
465                 470                 475                 480
Pro Leu Gln Arg Gln Pro Thr Met Cys Ser Ile Asp Glu Leu Thr Asp
                485                 490                 495
Asp Leu Ala Ile Ser Arg Ser Arg Leu Lys Leu Thr Arg Ile Thr Arg
            500                 505                 510
Pro Pro Gly Gly Val Thr Gly Gly Val Ala Gly Gly Ser Thr Thr Gly
        515                 520                 525
Ala Ala Gly Ser Gly Gly Val Ser Gly Asp Glu Ser Ser Gly Lys Val
    530                 535                 540
Arg Lys Ala Lys Val Lys Val Leu Lys Ser Ser Ser Pro Phe Lys Gly
545                 550                 555                 560
Leu Arg Thr Lys Phe Asn Trp Arg Ala Arg Arg Lys Gly Ser His Lys
                565                 570                 575
Pro His Glu Lys Gly Ala Thr Val Asn Gly Gly Asp Thr Glu Glu Arg
            580                 585                 590
Ala Ala Phe
        595


<210> 196
<211> 9889
<212> DNA
<213> Drosophila

<400> 196
taattgctgc cgggtaattg caactgagct caggtcagcg agtttgctaa ccatggcgtc 60
attatactgt tgcactataa aatcattcga cccaagtccg attaacttaa gtgggtctta 120
gaccaaaaga tgtgtatttc gaatacttct caaacgttat ttccttaagt caaatcattt 180
caagaactgg gagagcaaat aaatatgcga tttcctcgac ttatagttac ttgtgtgaaa 240
tttgcgcgct tgtgcaaatt cgaatgtttg caacttaatt ggactgcata atagattggg 300
caaaaggatt ttataacaat tccagtgtcc agcccatcgt actgatgtgg ctaattgttt 360
tgctgacttg cacctaggca aagcaatttc ccttgaaaac aatcaaattc aaaacaatca 420
agtaaacaag acagctgagt gaaagacttt ggcttcatgg gtgaacatat ttgtcggttg 480
aaggttgagt ttttgtttaa gatgattgca catggcacac aggttgattt taatctaaat 540
ttaatttgag catttgagca tttttctgct gaaccgggaa aacccacatt tgcttaccta 600
gaaaaaccgg ttttcagaga aagttgcgag taatacatat ttaatacgat ttcaagtcct 660
gtgtgtacaa cgttttatct aagaaaacca gaaaaccaga agctttttag cttagacagc 720
tttaatattt aattacagtg ttcacaaaaa aaaaaaacga aattcccctt ttctttgctt 780
caaatttaat tagacatgta tctattgttt tttttttcc acttgcagct tatttatcca 840
acctaagcca actaaatgat gcaatgctgg tgacatataa aacgaaagtg gaaacggagt 900
tcattcctta agtaaccaat ttaagatgtt gtgctcccaa gatccctaaa ccaccgtcca 960
aaccccttcga ataacttatt taagccaaga tatagacaaa atggaggacg agtggggctc 1020
ctttgatcgt ctgcccagtg ttccgagtgc ctcgatggat ttggagacgg aaaacgaggt 1080
ggtcagcaat tggtccacac tggccaactt cacgcgactt gtggctggtg ccgctcctga 1140
aatcgtcaac tatacgctca acatgatcga cgtgggtgtg ggcatggcca cggatatatc 1200
caatttgagc gtgagcacca cgcccctgcc cgcctacgca atctccaata gctcctcact 1260
ggcgcacacc aatagtcgcc atgaagcacc accgatggcg gaacaggttc cggagcacgt 1320
gatggatcac gcacctcaac tatcccgatc cgggttgctg aaagtgtatg tcctggcggt 1380
aatggcactg ttctctctgc tgggcaacct gctgaccatc tggaatatct acaaaacccg 1440
catctcaaga agaaactcac ggcacacgtg gagtgctatc tactcactga tgttccatct 1500
gtccatcgcc gatgtcctgg tcacctggtt ctgcattatc ggggaggccg cgtggtgcta 1560
caccgtccag tggctggcca atgagctcac ctgcaagctg gtgaagctct tccagatgtt 1620
tagcctctac ctgagcacct atgtcctggt cctcatcgga gtggaccgct ggatagcggt 1680
caagtatccg atgaagtcgc tcaacatggc caagaggtgt cataggctac ttggcggtac 1740
ttacatcctg tcgctggtgc tcagcttgcc acaggtaagt tttgatacac tgaaaataag 1800
acctttaaag ttcttataaa atcatttgcc tagaacaaac agataagata gttacacaaa 1860
```

taagcaagat tgtaataata taaaagtcat ttttggccta taaatggtct aaaaatattt 1920
gtgggatatg ttagattcaa tctttagata atttaaatga aaaaaagtta aaggtgccct 1980
ttgattagga aatcagtaaa agctttaaga tctcctctta atttatgtac cctttcggaa 2040
aatttctgtt cgacgtaaat gcaaaaagcc ttgagggtac ccgtaaaaac attattctat 2100
ttaaaaagat ttagccaaag ggcaaattta aagcgtctgc atctgaaaaa atggcccttc 2160
cagtgaagtg cgaagaagac acgcttatga aaataatcca gtcacttttg gtggggcgtt 2220
ccattacgct aagttaaagc tgcgacttag ccgtggtaat attcacctgc cactgcccct 2280
gcctgtgccc cctactcctg gcactccatt tgatccttac gctcgcgctg atgaagttga 2340
aataatatgt ttactttata acacgtatac gccccgtggc acagtgcaaa tagtcttgag 2400
taggcgccag aatgtgtgca gacattaaaa atgattgcct ttcgcacaat aagtgtacac 2460
taaatgtggc tttaggccgc tgccacgccc acacacacgc gtgctgatgg gcgtaggaat 2520
gggcataggg ttgggcgtat ggggacaccc aatcgaattc cggtcgtccg accgacaaag 2580
ggttttttc ttctggtttt tcctctcggc ttttcgcttg gggacgggcg accaaataaa 2640
aaaaaaaaaa ggtccgaaca agtgtgtctt ctgcctgacg acagacatta ttcatattcc 2700
cgatttcggt tctttgtgcg cagataggca cactgggtga tatattcctt tgccctggca 2760
gtcagctggc ggataagcct gaatgggact gcattaatgg gcgtcactcg ctttgttcta 2820
caactctaga acgactcggc ttcagatttc cggcaccctg attgtataat atatgccgcg 2880
tgtctccaca tcccgatttc atactcgatt cgcctggcct ttccttgatt tatggccccg 2940
ggtgaggcac ttcacttctc gaattcagtt aatctttagc tggagtgcta cgagctcctg 3000
ctcgtgtttt tccttaaagt ggaaattcgt ggacacttca gtttagggtg tcactaaggt 3060
aggtaattag cacgaaatct cacttcaaag ttgccaaggg ttagtaatat agttagattg 3120
aaaagtgact gcaaactttc cactgctctt tatgtgaccc gggtcataaa tatcgttcgc 3180
agaatttggt tagccataag tggggacaag tgggtagaaa ggtcaaacct tctgtataaa 3240
aagatttgtc acctgagaga tgccaaaaat agatgaaaaa gaaccaaagg gaaatttcct 3300
cctaagtgga tggtaagtgg atttcccgca cgcaagttat gcaactagca cactttcccc 3360
actgctcgcc cgcaaatgta tgctaaaatg tagctcattt aattagggag cccagcaacc 3420
cgacgccatt tcctttttatc tgaataatgt attgcgtgga acaataatta cagacgtacc 3480
gtacaaaaat cgaggcgaga ttatatttgc tcactttggg gcaaaatgtg tgtgtggttt 3540
ttctgggtgc gaaggcatgc gcgagtcagt ttgttaaatc atctgacttg tttaaatgcc 3600
gcaaaaaatc tactgagctt aataacattt tgatgttttt ttccaatcga tgtcaagcgg 3660
aaaatttggc cgttttgttg ctgggaattt gggccgactt ttggtccagc cgacccataa 3720
tggaaaggac ataatgtatt ggttttggat ttttaaaata gaaccagact ttgatgaact 3780
gtccagcttg ggtgtaaata tttgatgacg aacgtcacaa attacgatga tatgacagaa 3840
gtcatttatg gatggccact gacacccaat taaaagaaac tttgcaatca ttaagccaat 3900
ttttatggtc cacttttcga catctcgctg agatcccggg aaactttctt catttcgctt 3960
ctactcgaat atgcgacata aaatttatgc cgcttacaga cgtgacggtc gaaacgttcc 4020
cgccggaaac gtaggaatca tctcaaacaa ttttccggat aactttttcc atgtcacacg 4080
ttcatggtgc aggcaaccaa acagccaacc aaccaaccaa gtggcacatg tattggtaag 4140
gagccacatc gaatccaata tgagacgcgc cgctgaactc gtcggtttca aataaaggag 4200
cgaaatgaaa attcacttcg actgccaata aacaacaaat gttttcaaca gcagcattta 4260
tttggctcca cttctatcga ccccaactcc tccgcgtcct tcgaacccaa ggaaaaaaga 4320
aaaaaagttt agtaccacac gcacacaacc acacagaaga atcctgccac aataactcga 4380
agtgtgagcc caaaaatgtg gcacaatgaa aacaggccgc cataatacag tcaaacacaa 4440
agttatccac aggtcctccg cagagtatta ttgcgaattg aatttcaccc acatatcatc 4500
aaattcagca gccttaggtg gcataattaa aatttcgtag aattgtttac aaattggagg 4560
gaactagtca atattacaag agtcccttgg cgatttcacc ttcagtgtgc gcgggctttt 4620
gataatcgat aggggtcctt aaatgcttag ctttaagtga attaaaattt gaatttagtt 4680
tcagaatatt tacagcagaa tttgaatcaa taattgcaat ttatatgtgc aagcattggc 4740
gaataaagta aataaccgga tttaattatt aggtgacaaa atcgtaacat ttaaatatat 4800
ttacctgaat aacaattgga aataagtgca aaaacagcaa aaagtataat aaagtttaga 4860
taaaaaatga aatccattta atattttaat ttgggctcat ctggattcat attttaggtg 4920
acaaaattaa ttattataat tattaaaaca aatgagcagg taacaagtca gtaagctcta 4980
acaatccttg ttaaatctaa cttaaacgta gctgagtgtc caatttaaat ggtcaattta 5040
atcgtgcctt taagccaaac aaatatgtat attttaacaa tctgatagtt ttacaaacgt 5100
cttaaacgat tgccaagcac ctgaccgatg tgtgaaaact acgaacttga gcaatacaat 5160
ttggttagaa cggaagttct tgaagttgtt ccctcacaat gaatcaatct tgttttattt 5220
tactgtcgca agtttttgtc taaagttttt cgttgcaact tgtctaactt tgccgcaaat 5280
agtgctgaag gcttgcttgt gtgtggcatg gttttagcgt ggagcgccct ttaaactttt 5340
cgagcatttt taatgaatgg caacatggaa ataatgaaca aacacttgaa aagttgaaat 5400
aatttaagaa agaatattga aagccacaat tatccttgct ttcttttact agtcatgacc 5460

tataaaaata tatttttcta ctataaaaag ccactggcac tttacttcgg tttcggcgga 5520
aggcaatggc ttggaagtta taggcgaaat taataaaaat caaccattag acgattatcc 5580
gcttctcgtt atttctgcgc ggggggtttt tgtgtagttt ttactacgcc actcggcgga 5640
gcgcggaaag ttttcccatt tcctccattt gggtcagctg ttgacacagc ggccgccaga 5700
aaacccataa aaagcgcaaa taaaaatact tttagcgtat tagttatgcg tagacacatc 5760
catcgttggt gtatttaaga tgctacactt cctttatcag ttatggccac acattttcct 5820
gctcgaaaag ctttttcatg gtacgagaat ttccaccttt ctacgtggca ctctgctaaa 5880
ctcggtaaga cttgaacaat ggagcagtag caactgcaga atgtatacat tttccaaaaa 5940
tgttccatca atttggaaat gatttcagcc aaacaagaaa atcaagcaaa agcacttttc 6000
acccataacc aagccgcttg ccacccacaa tggaaagtag atcaaagcca caacgttttg 6060
agtcagcagt ttgttaccat ggaccgcatg tgggtatttc catgtaaaca tgcctgtact 6120
tagtccttgg tgcaatcagg acatggatgg tttgctaata aaaatcttat tacggcacca 6180
tcttatttac acaacacggt agagatatat ggttttaaaa atttccattt tcctgcttaa 6240
aacaggcaat aaattgatag gaaggtaaat atccttaaga aacacttgta taaagttcaa 6300
gttcaatcca tgtgtcttaa cctttaattg tttactatgt tgcatgttcg ttccagttct 6360
tcatcttcca tgtagcgcgt ggcccattcg tggaggagtt ttaccagtgc gtcacccacg 6420
gattctacac ggcggattgg caggagcaga tgtacgccac cttcacgctg gtcttcacct 6480
tcctgctgcc gctgtgcatc ctgtttggca cctacatgtc caccttccgc accatttcca 6540
gtgagttttc agcgagaagg agtttacatt cccgtactaa cctgataacc tgataatcgt 6600
gcaggcagcg aaaagatgtt ccagggatca aagttggcca actactcaac ggccaaattg 6660
cccacacaga cgaatcgcca gaggctgata cacaaggcca agatgaagtc gcttcgcata 6720
tccgtggtga tcatcatagc gtttctcatc tgctggacgc cctactacgt catgatgatt 6780
atgttcatgt tcctcaatcc ggacaaaagg gtgagtttcc gcccgttccc ctgccgaatt 6840
tgaagctgaa tgacaagcag atatcgattc ctgataccaa aaggctcatc aatcagcagt 6900
cgatttgtca gtttgtagcg cacggtttgg gcgaatcctt tgtcctgttc cctaccatat 6960
ttacgacttg ggtcatggt tgtacgagta aaataaaaac gagaatacga cacgacctag 7020
tcgcagctgg caaatccgtt tggagctata cccactgcca taaaaatgat ttcttatgct 7080
caatttttaa ttaatcgtta cgccaccga ccgaccacac ggtcgcaatt tttccttaac 7140
ggcgtaccca actttgatga cttgccagag atggtccgta aaagctggcc caagttgcaa 7200
gattacggcg aggcaatccg ggcagaagga acacaaggtc gaattgctac aaaccgagca 7260
ttctgtacag taggggaaaa tgaatgagta gcaaaatatt tccaaagttt atttgtttcg 7320
atttaaattt taaaaatgct ataattgaca aaatagaatc atttttgctg tcactatttt 7380
taagtgaatg ggatatttgt ccgccgatta acctgaattt catatcctaa tgaattggat 7440
tgcacattca taatggcatt tccgctgaaa ctctttctcg ctgtggaaat cccagcgatg 7500
gcagtgtgca aaataattaa tttggaagcc aaatttttag cgcagtgtat ccataagcga 7560
cctgagcccc aattcgaccg caggaacgcc cacaatgcgc ccacgcgaat gtagaggggа 7620
aaaggtgaaa ggtgtgcgac aagggcaagg aaattaattt gaatttcacg gattgcttgc 7680
agctgggcga cgatctgcag gacgccatct tcttcttcgg catgtcaaac agcctggtca 7740
acccactcat ctacggtgcc ttccacctgt gtcctggcaa agggggcaag tcgagcggcg 7800
ggggcggcaa caacaacgcc tacagcttga acaggtgagt ggtgagccaa cagaataagt 7860
taattctatt aaccaccaca tttcgtactt gtttttgcgc ttttatcgtg tcaaaaaata 7920
gttacttttc actcaagcaa tacactttta agacagccaa gcaagctttt gaagtctgaa 7980
ttgctgcttt tctgatcaaa gagtttttat tcttggtaat atttcacatt taagagtata 8040
aatgatagtg taaatttaaa taaaatagta gaaagttttc cctaaagatc gatgtataac 8100
tattaataca tattttactt aatctagtgt tgctttttaa tgacagtttt tgctgtgctt 8160
aaatttttac aaaataaatc ttaaatcctt gcaaagagtt ttcttttatg caaatggaaa 8220
acattgtttc atcgcttgtt aatatcattt actttattta tttattattc taaacagctg 8280
ccaaagcagc gaactatctc aggatcagat atgatttgct atgattttgt cacacttcct 8340
atgtttgttt tgataacgca aaaagcacac acaccgtgat tggcaataaa ctgataaatt 8400
cttttccgtc tggtttttat cacctcaccg catcgtccat tccccсgtcc actttccatc 8460
cgcagggcg actcgcagcg cactccatcc atgctaacgg cggtgacgca ggtggacggc 8520
acaggtggca gttcccgcca gatgcgggcc ttccgccagc agagctacta ccgcagctcc 8580
tccaacggca cagccggacc gggtgcagct ccctttaagg agcaggttgg cctgctgcac 8640
gtgggtcccg gcaatgggac gcccggtggc tccgtctcta gcggcgccac gccgcagttg 8700
atccgtaagg gatcggctct gttggcccga caacccagct gtctgaggga gcaggagcac 8760
cagcagcgtt tgctgctgca cgagaagccc tcgaccctgg tgctcagcta cgacagccag 8820
cggggcggag tgggcgtggg cgtggccagc ggtctgctgg acaacaacga gcgagtgtcg 8880
agcgtgtgag agcagctggc gatggcggta gcggcggcgg cggtggcggt ggatgcaccg 8940
gacacagtag aagcagcagc agcaacagca tcagccgatg cggatgcgaa tgtggcgatg 9000
gaggacgcct gctcctgctg ccattgcctg agcagggagc tggtgcggcg gcagttggag 9060

ctccaagtgc ttccagtgca gtgaggcgac tgccagtgga tttgatccgg ctggctgagt 9120
ggccggagct ggggattcca gcgattccct cggtcaacgt ggggagtggc gtcgatgatt 9180
agatcaccgg ggcgtgcaac atgtgcttag atagtgctaa tcgagagtta agtcttcaaa 9240
cacaatgggt gtgcttagaa gtagcgtagg attagagcta gtccgtacga agtggttgcc 9300
ctacactggg agaaactcgt ggtcactttg catgttctct gaatgttcta ctaatacttg 9360
acttcaacaa ggatacaaat ataagagata atatctatgt acatcccata aaaccataaa 9420
aaattgcaat taaatctggt tgtgaaaaaa atagatttaa cttctgaatg tttcaaatgt 9480
tgccaattcg ccagcgtgta atccgatata agcaacaata attacaatta taatcaatgt 9540
gcaatcagag tcagttacaa tttacaatta tatgactgaa taattggact atgaatatat 9600
acgacttcaa atgagtttct ctcattagca gggcaacttt gtaatgttaa ttaacagcgt 9660
ttttcagttt ttcattttgt tcccccagcc gagtaactca aaaatcgatt ctataattgt 9720
gtaaaataga aatcatctaa aaacagggta aatctaacta ttgtaaatac atttactcat 9780
tattttgaga agaatattgt aaaaaaaaaa gattgtaaag tttaaaataa atatgagttc 9840
atgttcaacg acacaaggaa ttattttgat taatttaaag tactaaaaa 9889

<210> 197
<211> 1644
<212> DNA
<213> Drosophila

<400> 197
atggaggacg agtggggctc ctttgatcgt ctgcccagtg ttccgagtgc ctcgatggat 60
ttggagacgg aaaacgaggt ggtcagcaat tggtccacac tggccaactt cacgcgactt 120
gtggctggtg ccgctcctga aatcgtcaac tatacgctca acatgatcga cgtgggtgtg 180
ggcatggcca cggatatatc caatttgagc gttccggagc acgtgatgga tcacgcacct 240
caactatccc gatccgggtt gctgaaagtg tatgtcctgg cggtaatggc actgttctct 300
ctgctgggca acctgctgac catctggaat atctacaaaa cccgcatctc aagaagaaac 360
tcacggcaca cgtggagtgc tatctactca ctgatgttcc atctgtccat cgccgatgtc 420
ctggtcacct ggttctgcat tatcggggag gccgcgtggt gctacaccgt ccagtggctg 480
gccaatgagc tcacctgcaa gctggtgaag ctcttccaga tgtttagcct ctacctgagc 540
acctatgtcc tggtcctcat cggagtggac cgctggatag cggtcaagta tccgatgaag 600
tcgctcaaca tggccaagag gtgtcatagg ctacttggcg gtacttacat cctgtcgctg 660
gtgctcagct tgccacagtt cttcatcttc catgtagcgc gtggcccatt cgtggaggag 720
ttttaccagt gcgtcaccca cggattctac acggcggatt ggcaggagca gatgtacgcc 780
accttcacgc tggtcttcac cttcctgctg ccgctgtgca tcctgtttgg cacctacatg 840
tccaccttcc gcaccatttc cagcagcgaa aagatgttcc agggatcaaa gttggccaac 900
tactcaacgg ccaaattgcc cacacagacg aatcgccaga ggctgataca caaggccaag 960
atgaagtcgc ttcgcatatc cgtggtgatc atcatagcgt ttctcatctg ctggacgccc 1020
tactacgtca tgatgattat gttcatgttc ctcaatccgg acaaaaggct gggcgacgat 1080
ctgcaggacg ccatcttctt cttcggcatg tcaaacagcc tggtcaaccc actcatctac 1140
ggtgccttcc acctgtgtcc tggcaaaggg ggcaagtcga gcggcggggg cggcaacaac 1200
aacgcctaca gcttgaacag gggcgactcg cagcgcactc catccatgct aacggcggtg 1260
acgcaggtgg acggcacagg tggcagttcc cgccagatgc gggccttccg ccagcagagc 1320
tactaccgca gctcctccaa cggcacagcc ggaccgggtg cagctccctt taaggagcag 1380
gttggcctgc tgcacgtggg tcccggcaat gggacgcccg gtggctccgt ctctagcggc 1440
gccacgccgc agttgatccg taagggatcg gctctgttgg cccgacaacc cagctgtctg 1500
agggagcagg agcaccagca gcgtttgctg ctgcacgaga agccctcgac cctggtgctc 1560
agctacgaca gccagcgggg cggagtgggc gtgggcgtgg ccagcggtct gctggacaac 1620
aacgagcgag tgtcgagcgt gtga 1644

<210> 198
<211> 547
<212> PRT
<213> Drosophila

<400> 198
Met Glu Asp Glu Trp Gly Ser Phe Asp Arg Leu Pro Ser Val Pro Ser
1 5 10 15

```
Ala Ser Met Asp Leu Glu Thr Glu Asn Glu Val Val Ser Asn Trp Ser
            20                  25                  30
Thr Leu Ala Asn Phe Thr Arg Leu Val Ala Gly Ala Ala Pro Glu Ile
        35                  40                  45
Val Asn Tyr Thr Leu Asn Met Ile Asp Val Gly Val Gly Met Ala Thr
    50                  55                  60
Asp Ile Ser Asn Leu Ser Val Pro Glu His Val Met Asp His Ala Pro
65                  70                  75                  80
Gln Leu Ser Arg Ser Gly Leu Leu Lys Val Tyr Val Leu Ala Val Met
                85                  90                  95
Ala Leu Phe Ser Leu Leu Gly Asn Leu Leu Thr Ile Trp Asn Ile Tyr
            100                 105                 110
Lys Thr Arg Ile Ser Arg Arg Asn Ser Arg His Thr Trp Ser Ala Ile
        115                 120                 125
Tyr Ser Leu Met Phe His Leu Ser Ile Ala Asp Val Leu Val Thr Trp
    130                 135                 140
Phe Cys Ile Ile Gly Glu Ala Ala Trp Cys Tyr Thr Val Gln Trp Leu
145                 150                 155                 160
Ala Asn Glu Leu Thr Cys Lys Leu Val Lys Leu Phe Gln Met Phe Ser
                165                 170                 175
Leu Tyr Leu Ser Thr Tyr Val Leu Val Leu Ile Gly Val Asp Arg Trp
            180                 185                 190
Ile Ala Val Lys Tyr Pro Met Lys Ser Leu Asn Met Ala Lys Arg Cys
        195                 200                 205
His Arg Leu Leu Gly Gly Thr Tyr Ile Leu Ser Leu Val Leu Ser Leu
    210                 215                 220
Pro Gln Phe Phe Ile Phe His Val Ala Arg Gly Pro Phe Val Glu Glu
225                 230                 235                 240
Phe Tyr Gln Cys Val Thr His Gly Phe Tyr Thr Ala Asp Trp Gln Glu
                245                 250                 255
Gln Met Tyr Ala Thr Phe Thr Leu Val Phe Thr Phe Leu Leu Pro Leu
            260                 265                 270
Cys Ile Leu Phe Gly Thr Tyr Met Ser Thr Phe Arg Thr Ile Ser Ser
        275                 280                 285
Ser Glu Lys Met Phe Gln Gly Ser Lys Leu Ala Asn Tyr Ser Thr Ala
    290                 295                 300
Lys Leu Pro Thr Gln Thr Asn Arg Gln Arg Leu Ile His Lys Ala Lys
305                 310                 315                 320
Met Lys Ser Leu Arg Ile Ser Val Val Ile Ile Ile Ala Phe Leu Ile
                325                 330                 335
Cys Trp Thr Pro Tyr Tyr Val Met Met Ile Met Phe Met Phe Leu Asn
            340                 345                 350
Pro Asp Lys Arg Leu Gly Asp Asp Leu Gln Asp Ala Ile Phe Phe Phe
        355                 360                 365
Gly Met Ser Asn Ser Leu Val Asn Pro Leu Ile Tyr Gly Ala Phe His
    370                 375                 380
Leu Cys Pro Gly Lys Gly Gly Lys Ser Ser Gly Gly Gly Gly Asn Asn
385                 390                 395                 400
Asn Ala Tyr Ser Leu Asn Arg Gly Asp Ser Gln Arg Thr Pro Ser Met
                405                 410                 415
Leu Thr Ala Val Thr Gln Val Asp Gly Thr Gly Gly Ser Ser Arg Gln
            420                 425                 430
Met Arg Ala Phe Arg Gln Gln Ser Tyr Tyr Arg Ser Ser Ser Asn Gly
        435                 440                 445
Thr Ala Gly Pro Gly Ala Ala Pro Phe Lys Glu Gln Val Gly Leu Leu
    450                 455                 460
His Val Gly Pro Gly Asn Gly Thr Pro Gly Gly Ser Val Ser Ser Gly
465                 470                 475                 480
Ala Thr Pro Gln Leu Ile Arg Lys Gly Ser Ala Leu Leu Ala Arg Gln
                485                 490                 495
```

```
Pro Ser Cys Leu Arg Glu Gln Glu His Gln Gln Arg Leu Leu Leu His
            500                 505                 510
Glu Lys Pro Ser Thr Leu Val Leu Ser Tyr Asp Ser Gln Arg Gly Gly
        515                 520                 525
Val Gly Val Gly Val Ala Ser Gly Leu Leu Asp Asn Asn Glu Arg Val
    530                 535                 540
Ser Ser Val
545
```

Celera Sequence No. : 142000012783465
CAAAAACTAAAACAACAAAAAGCCGGCAAGCAATGGTATACAAAATAAATAAAAAAAAATATGACAAGAGAAAATCCACAAAAAAAATGAAATAAAATTGAATGAAAAATTTA
CCAAAAACCAAGTTGTTTTATTCCAAGTGTATGTAGTTGGCAATCAAACGATTCAACTGCAGTTAACATGATTTATTCGCTACGTAAGTGTGCTTGGCTAAGTTGAATAAA
TATCATTAATATCTGTCAGCTACATTAAAGGTTCAAACATTACATATATATATATCTATATATTTATCTTAACATTATTAAATGTTTAAGGGCTTAAAAAAAAATATCCTTA
GGTACACGGGTACGTATGTATTATTCTGCGTCTCTTTCTATATCCTCTCAGGATTTGAATATATTTTGTATTATAAAATATTACGCATAAAATCATCGGGGGGAACCTGTTA
ATTCTTGGCATTTGATTCCCGCATTTTCTCTGGATATGGCATAGATATTTTTCTAGGGTTAGCTGTGTGAAACCTAAATATTTTTAAGTTAAATTTAAACAACGATTTATAC
GGCTGACACATTTTAAATTTTTCTGGGTTCAAGCTTAGATTTTGTTTACATGTTTTTTTCTTAATGCTTAACACTATGCTTTATATACTTTTTCAAAATATATTTTAAAACT
TGTTCAGTAAAAAGTTTTAAAGTGATAATTCTTAACAGAGTTTTTTTAGAAACTTTTTTTTTAGAAAGATCAAGCATACAATTATTTGTTTTTATCAATAATTCATATGACT
AGAACCTTTCTTTTCTAAATGCATTTCGGGCATAGTATCCCAGGATAAAAAATATAACAGGAACATTTTAAATAATTGAAATCAGTATGCATTTTAACAAAATTAAGTTATA
AAAAAAAATATTTTGCTATCATGAAATAAGTGTGCCGGACTGGTCTTCAAAATCTTTGCTCGGCGATGCGGTCAATGTTACGATCGACCCTCGCGCGGTACGCCTTAAACGC
CCAACTCTCGGCTGGGCGTGGTCGTGGCCATCTGTCCAGTATTACTCCCGGCCGGCGGCGGGGTCACTAAACTGGGCACCAACTTGTTCAGCAGCGTGCGAATGGCAAAGGT
GACATCGTGATTGGCGAAGCAGAACAGGAAGGAGACCACAAAGCCCTGCAGGCTGACCAACACCACCGATAGCATCTGATAGAAGTGGTCCAGCTGGGTGCCGGCATCCGGA
CGGTAGGGCAGCAGGAAGTGTTGCAGTCCGAACAGCGGTACCAGAATAATGGTGGCCCGAACCGCCTTCCGGATGGCCAACGGTGCCGGCTGGGCGGACTGGGGATGTAACT
TTCGCACGATCACTCGCAGTACGTTGATCAGAAAAACTGTTGGGTAAAATGCATATAGCTGAGCAATATTGTACTTTTAAAGAAAATTATGCATATTACTAAAGCTGGCCAG
CAGCGAAAGAGTTATGGGCACCGAGAAGATCCACAAGTAGAGACTGTCAGTTATCCAGCAGCTGCAATAATGTGAAAAGAAATGTTACCATGGATATAGCGTAGAATGTAAT
CTCTTAATAAATGCAATGCAAAAACCAAATCAATCAGGATTATGTCTTGTGTAAAATTGTGAAAATAACTTACTGCTTATTGTCCGGACTACTGAAATGCCTGGCCAGACCG
TAGACGATGGCTATGGGTATGGGTGAAAACCAGCTGATGACAATAAACCAGCGCATCACAATCGTGTCCTTGACAAAGACCTGGCGAATGGAAAGAAGAAGGGGCGTCAGGA
TCTGAACGGACATTCCTGACATCTCAACTTGACGCAGACGTAAAGCAGTTGGCCAACGTTGCGATTTGCTCTCAGACAAATTGCTGGGGGTGAACCATTGCCATGTCAGCGG
CTGCAGGAAATCACATCTGCCATGCACAGAAGTCGCCCATAAAGAATGCAGACTTGTGCAGGCAATTGCCAGATTGCCAGACGAGTAAGGATAACTCACCACAACGAGCACC
AGGTGCAGGTGCAGGCCCTCGCAAAACATCCAGAAATAGTTGACCAACATAAAGTAGTGCACCACCAGATGCAGCCCAATGCACCACAGCTGAAAGTTAGTTATGTTTATGG
TTTTGTATAATTGCCTAGCTACTTGTAGGGCTATGTTATTTTCATAGTTACTTTTCTCAGTGCATCTTACCGGATTTTCGGCTATGGTTTCGGAACGCTCCACCACCAGACG
ATACCATAATATCCAGGCCACACAGGTGCAGGCCAGCGAGGCGAACAGATGTACATGGATGCGGATGCGGGTGCAGCGCAGGGATCTGCAAACGGAATGGAACGGAACGGAA
AGAAACGAAAAGGGATATCCGGCCGTCAGTTACCGGAGGGAAAGGAACCGTGTGGGTCGTAGACAAACACTTACTTGAAGCCCAAGAAAATAATAATCGATATGAGCAGGGC
CAGCAGGGAGAGGGCATAACCCTTCACGTACAGCTCGTTGATGAACTGGCGAAACTGCAGTGGCAACTTTCTTGATTAATGCCCATGTACCACCCGCTTGATGCCGCCGAAG
CACTTACCTCCAGGTCCTCATAGTCCACGCAGTTGGTGTAATTGGACCAGGTCTGATTGCTAACCGGATGACGGTACCAGGAGCCGTTCTCCAGGCAGCTATGAATGCGAAA
TGCTTACGATAATTGTGACTGGCTAAGTGCGGACTGATGGGTAAAGGGGGTATTTCTAACAAAGCGCCTGACAATGCACAAGTTACCCAACTTTACCAGGAGCCAGCTGAGC
TCAGTTGGCGTGTAAAGTGAGATGAACGCTATCCATTAAATTGGAGATACATACATGTAGTGCAAGTTTTAGAGTGCTGCTACAGTACACTTAGGGAGATTATCTCTGCAAG
AGTTGAAGAAGGTTTTATTTTTCTTTCATTTTCCATCATCTTTTTCTGAGATATGCTGCGAGTTTATGTAATAGACACGAATGTAAACTCTATAGGATTTTCCCTGTTTGTA
GCGCACTCATACATAAACATTGTATTTTATGACACACAGAACACATTAAAGCAAATGCCGAATATGGCAAGATATTTTCGAATTTATGCTCAAACGAATATTACATATTGCA
GAGCTTATTTGACCATGCGAAATCCTGAACGGAATGAAAATCTATCCTATTCGTGGTGGCGTAATATACAGGCATATTGGGCATTTAACTCACGTCTTGTGGGCCAGAAATT
TTCTGTTGAATCCTTCGACGAAATCGGGACAATATTGACTCAGCACAGTGCCGGCCGGAGTCCTTGGCCAGCAGAGATATCCGTCAAAGTTCAGGGGGCAAAAGAGTCGGTT
CTCCTTCTGATTTGTGGCCATCGTGGCACTGACGACGGTGGCCATCAGCTCCTCCTTCTCCGCCAGCTCCTCGTCCACGTTTCGCATGTCAATAACGCCCTCGTTCATCGCG
GCATTCTCCAGATATCGCGGCACTGGGCTAAAGTTCTCCGGCAAATCGGAATCGTAGTCTGTGCACGGCGACAAGTGCCATTAAAATGGGTTTACTTTTTAATGTATAAAAA
TATAGCATTTTTAATAATAGTATCTATGGGCTGCATAACTGAAAGTGATGTAATTCAGTTTTCAAGTTCAATTCAATTGCATGCTTGCTTTACATATGTAAAATAAAGTGGT
ATTTCAAATGTACATAACTTGCTTAAGATAGTTAATAGATCATCGTTCTTTATAGCAGGGTATAATTCATAAATTCTGGTTGAAAATATATGTACATTTGGTAAAGGTGTTG
AGTGCTATTTAAATTTATAAATTTGTCCTGAGATGCCAATGATTTGCATTGAAACGAAAACCCATGTTTAAATTTGCATCCCCAAGAAGTAAGTCAATGATTTTGTATATAT
TTAAATACATTACTTTGAAATGCATTAAATGAAATCAGATGGGCTGTAATTTAATTTAATTCTCTTCTCTTTCTTATCTGTCTTATGCCTCTGCTATTTTTTCGCGGCAAAT
CACGTATACGCCTGTTGGCCGCCACTCACCTGTTGCTGGGCCAAGTGAAAATGATGGGTCGTCCGTATCGTATGTGACATTCTGGTATCGGAACACGCACTCCGCATACAAA
TGCTTTAAAAATATGCGCAGATTATCCTGCGTTTGGGCCTGCGATGCCACATCCACATTTTGCGACTCCGAATCTGAAGAGGTTGTCGCCATTGTGGGTTATTCAACGTGAC
TAACTCTTGTCAACGGCTAATGCACTAAAAAAAAATATAATAATATAAATATAATTCAAAGAATTACGGGATAAATAGGATAATCAACTGCAATAATTTATCAGAGGCAATT
TGAAAGGTTTTAAACAAGTGTTGCTTACCCATAGAAACTTGTATGAGCTAGTTTTTAAAATTTATTCTGATTTAAATATATTTTCTTATATTTAACATTAGTTCTCACTGTA
TTTCGCCATAGCGAAGCGAACTTGAGTCCCTTGTGGCCGTCCATTCGTGTGTGCATTTCTAATTTGGTTAGGCTTCAAAGAGTTGCCTGAGTTGGGTTGATATTTGTTGTGG
GGTCCTGTCTGCGGGGAAATGGTGCCAAAGGAAATGGAATATCCACACGGGGCACACACACACACTCAAAATATCGAAGCACCTGAGTGTCGTGTGAGCCGAGTCACCTACT
CGTCTCTCATATGTTGTAAATCAGTTTATTGTCTAATTAATGGCGGTGCTTAACTGTGTCGCTATTTCTGCTACCGTGGTTGGGCGAACTTAATTGCATTTCGTCGCACAGA
GAGACAAAGCAAATGCCACGGGTTAAAAACTCAAATGAAATGCACTTTGATCGTGGGGGCACACAAACACACAAACACATAAACACACTCACACACCCGCACCGAACACGTG
GCCCAAAAGGCGGGATGAGCTGCAAAGTTCTTGCCTAACTTGGTTATATTTTTAATCAAAACTCATTTCACTTGATTGAGTTTCGACAGAAGCCTGCACCAGGCACTGGAAA
ATCTAGCATGCACGATACATACAAAGTAAGATGATTTTCCCGCTCCCAAGCTAAACTTTAATTTGCGGCATTAATTTATTACACCGAACTGGGACAGTTTCCCATGATGGAT
ATGACGCCTGGCTAAACATTCTTGCATGTTATATATATTTTTGCAAGTGTCTCGGCCAGGCAATTAGCCGCAGGTAAGCGGAGC
(SEQ ID NO: 1)

Exon: 4236..4062
Exon: 3530..3230
Exon: 2674..2584
Exon: 2518..2427
Exon: 2325..1679
Exon: 1380..1001
Start ATG: 4236 (Reverse strand: CAT)

Transcript No. : CT4121
ATGGCGACAACCTCTTCAGATTCGGAGTCGCAAAATGTGGATGTGGCATCGCAGGCCCAAACGCAGGATAATCTGCGCATATTTTTAAAGCATTTGTATGCGGAGTGCGTGT
TCCGATACCAGAATGTCACATACGATACGGACGACCCATCATTTTCACTTGGCCCAGCAACAGACTACGATTCCGATTTGCCGGAGAACTTTAGCCCAGTGCCGCGATATCT
GGAGAATGCCGCGATGAACGAGGGCGTTATTGACATGCGAAACGTGGACGAGGAGCTGGCGGAGAAGGAGGAGCTGATGGCCACCGTCGTCAGTGCCACGATGGCCACAAAT
CAGAAGGAGAACCGACTCTTTTGCCCCCTGAACTTTGACGGATATCTCTGCTGGCCAAGGACTCCGGCCGGCACTGTGCTGAGTCAATATTGTCCCGATTTCGTCGAAGGAT
TCAACAGAAAATTTCTGGCCCACAAGACCTGCCTGGAGAACGGCTCCTGGTACCGTCATCCGGTTAGCAATCAGACCTGGTCCAATTACACCAACTGCGTGGACTATGAGGA
CCTGGAGTTTCGCCAGTTCATCAACGAGCTGTACGTGAAGGGTTATGCCCTCTCCCTGCTGGCCCTGCTCATATCGATTATTATTTTCTTGGGCTTCAAATCCCTGCGCTGC
ACCCGCATCCGCATCCATGTACATCTGTTCGCCTCGCTGGCCTGCACCTGTGTGGCCTGGATATTATGGTATCGTCTGGTGGTGGAGCGTTCCGAAACCATAGCCGAAAATC
CGGTAAGATGCACTGAGAAAAGTAACTATGAAAATAACATAGCCCTACAAGTAGCTAGGCAATTATACAAAACCATAAACATAACTAACTTTCAGCTGTGGTGCATTGGGCT
GCATCTGGTGGTGCACTACTTTATGTTGGTCAACTATTTCTGGATGTTTTGCGAGGGCCTGCACCTGCACCTGGTGCTCGTTGTGGTGAGTTATCCTTACTCGTCTGGCAAT
CTGGCAATTGCCTGCACAAGTCTGCATTCTTTATGGGCGACTTCTGTGCATGGCAGATGTGATTTCCTGCAGCCGCTGACATGGCAATGGTTCACCCCCAGCAATTTGTCTG
AGAGCAAATCGCAACGTTGGCCAACTGCTTTACGTCTGCGTCAAGTTGAGATGTCAGGAATGTCCGTTCAGATCCTGACGCCCCTTCTTCTTTCCATTCGCCAGGTCTTTGT
CAAGGACACGATTGTGATGCGCTGGTTTATTGTCATCAGCTGGTTTTCACCCATACCCATAGCCATCGTCTACGTTTTTCTGATCAACGTACTGCGAGTGATCGTGCGAAAG
TTACATCCCCAGTCCGCCCAGCCGGCACCGTTGGCCATCCGGAAGGCGGTTCGGGCCACCATTATTCTGGTACCGCTGTTCGGACTGCAACACTTCCTGCTGCCCTACCGTC
CGGATGCCGGCACCCAGCTGGACCACTTCTATCAGATGCTATCGGTGGTGTTGGTCAGCCTGCAGGGCTTTGTGGTCTCCTTCCTGTTCTGCTTCGCCAATCACGATGTCAC
CTTTGCCATTCGCACGCTGCTGAACAAGTTGCTGCCCAGTTTAGTGACCCCGCCGCCGGCCGGGAGTAATACTGGACAGATGGCCACGACCACGCCCAGCCGAGAGTTGGGC
GTTTAA
(SEQ ID NO: 2)

Start ATG: 1 (Reverse strand: CAT)

MATTSSDSESQNVDVASQAQTQDNLRIFLKHLYAECVFRYQNVTYDTDDPSFSLGPATDYDSDLPENFSPVPRYLENAAMNEGVIDMRNVDEELAEKEELMATVVSATMATN
QKENRLFCPLNFDGYLCWPRTPAGTVLSQYCPDFVEGFNRKFLAHKTCLENGSWYRHPVSNQTWSNYTNCVDYEDLEFRQFINELYVKGYALSLLALLISIIIFLGFKSLRC
TRIRIHVHLFASLACTCVAWILWYRLVVERSETIAENPVRCTEKSNYENNIALQVARQLYKTINITNFQLWCIGLHLVVHYFMLVNYFWMFCEGLHLHLVLVVVSYPYSSGN
LAIACTSLHSLWATSVHGRCDFLQPLTWQWFTPSNLSESKSQRWPTALRLRQVEMSGMSVQILTPLLLSIRQVFVKDTIVMRWFIVISWFSPIPIAIVYVFLINVLRVIVRK
LHPQSAQPAPLAIRKAVRATIILVPLFGLQHFLLPYRPDAGTQLDHFYQMLSVVLVSLQGFVVSFLFCFANHDVTFAIRTLLNKLLPSLVTPPPAGSNTGQMATTTPSRELG
V*
(SEQ ID NO: 3)

Name: Calcitonin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384694
GCCGTAAACTATGGCGATGCCCCACTGTGCCCGCACACAAATTCCGCCACCTCGCTCCGAAATCGTGTGTATTTGGTTAATAATAATCTAGTTGGGCATCGATTGTTCGG
CCAAAATGAAGCCACTATCCGTGGCGGAGTGTTAATTAATTTGCACTTAAGCGCAGCTGCTGGAAGCACTTGATAGAAGTGGGCATGCCGGTTCTCTTGGGAAGTTGGCT
TACTATAGTTTAGCGTTAAGGCGTTAACTTGTGGCCTTTAAAGTTCTTAGAGAAGACTTTCTGCAGCTGCTCAGGCTGTTCATTGCGGCATCTAGCATAAAATTTAGAGT
TTGAAAGTAAGTTTCTATGAAAAAAAAATACACGTATATATAAGTATAACAAGAAAAGGACCAAATCAAAGTTTATATACCATTGCTGTACTTATATATTTGAATATTAA
ATATACACAAATAAATATATACCGAATAAGAACACTTTATGATTAAAAGCAACAATTATTGATTAATCTACTTCTACCTAAACTGGTGTTGAAACTGGTATAATGAAAAA
CAACTAAGGTACATTTACTTCCATTTAATTTTCTATTGGCAGCTATATAATAAACAGTTCCGATCCAACGTGAACCGACCTATCACCATTTTTGATAAAGTTTTACGCAG
AAACGTAAAAATATCTAAAAGCATGCTATTTTGTATTTGTATTTTTGTATTTTGTAGAGAAAAGTTGTAGATCTATATTTAATTTTGCACACATGAGCATTTTAATCAGT
TTTGTATACATTATTTCGAACTCTTCACCAACAGAATAAAATGAAATTAAAATGGTCCGATGTCGGGGGCCAAACAAAATAAAGGCGTATTGAGTTGTGCCATGTTTCAG
GTGGAAAAACCAGCTGCAATATAATTGCTTTGAATAATGCGAATTTCGGGGGGAAGTTTCTTTTACCAGCACAATATAAACAAAACTGCTCAACTATTTCTCGATATTT
ACATTTTCAGGGTCCATCAGCAACATAGATGACGCGGAGTATCACTGCAATAAGACGCAGGTGCGCAAGGTCATCTCGGGGGTGGTGGGTGCCGCCTCCTCGGTCACATC
CATACAAGTTGCCAATTTATTGCGGTTATTCCGAATTCCCCAGGTGAGTTTCCATTGTGTGATCACTAACTGTTCGACTCAGTGCGTGCGAAGTCCAAAGGTCATGGCAT
AAAAGGGCAGTCCTTTTTGTTTTTAGTTCAGATTTTGCTTTATGCTCGCAACTACATGTATATGTGCTAGTTTGGCCAACAAAAGTTTCTCCTATTTCGTAATTAAAATT
TCCAGGAGCCACATGGTAACTGTTGCAATGGCTTGCAGGTCTTTCGGTTGGCTTTGCCGATTCCTAAGTGTCTGCACGTGCCACTAAATTAGTCGAGTCCCGCAGGCCAA
CTAGAGGAACATAGTGCTGCAGTTGCCAAGAAGTGTACAGTAGGGACTCTCAGCAATCATAACAATTAATGTAATTATTTTACTATATGTTCAATTTATCTTTAATATAC
AATTCCAAAATGATGATTCTTAGCTGCTAAAGAACTGACTTACATATTAATTTCCTACCAAAACCAATAACAAACTGGACAGATCGCTTCAGTTGAGTACTTCGACTATC
AGATACCCGTCACTTGCATTACCGCTTAAAAAGCTTGTATCTAGCATTGCTCGTATTTCAACTGCAAACGACATTGCAGTTTTTTGGGTGTTCCTCATGTAGTGACTAGA
GCGTGAAAAACGCCACAAACTAATATCCATTTCTGATTCAGGAAGGAATATAACATTTTCCGAGTGCCGGCATGAATCACGTGCAGGAAGTTGAAGGAGATTGAGTGCCC
GGTTCGGCTGCTCCAGCTACGTGAGCTGGAGCCAAAAGTCAACAACGCGACAATAACTAAGTAGCCAACTGCCAGATCCCTTCTCGATATATATATATAGTATGTACTCA
CTACTACTTGGATGCCATTAAACGCACACGAAGTGCCAGTCGAAGGAGCAGCAGCCGTTCCATGCAGGCCCAGCTATCGTCCCGAAAGGACCTCATAAATCAGCAATAAA
CATTTGATTGCATTTGGTTTGCACGCGTGGATTCTGCGCTGGCTTCGCATAAACAGCACTTTCGAGCAAATTTTGTGTGTTCTGCAATTCAAATTTCTGGTCGATAGCGA
CTTAGTGAGTTTCGTTTAGGCTTAGCGGAAATTGGCTGCAATGCAAAATGTTCGCAAGGACACAATTCAAATGACAATGCAGTTAAAAATTCCGTAAATGCAATATTTTA
TGGCTTATCGTATTTGGATAACAATGTTTTAAATGCTTGAACAAAGCATTGCTGTCTTTTTTGTTTAAAGTAAAATTTTTGCCGTCAGAAGTTTGTGCATTGAATTAAAT
GGTTATTCAAACATTACATTTTATTTATCTTGAAAAAAGTACAAAAATATGTATTAAACTCGTAGTATTATATATAAATTAAATATATTTTATTATTTTTTGTTTTTATG
GACTGAGTAATAGCACTTGAATTTTCCAACTAGTTTTTCACAATTTTTCGTACTTTAATAAAAGTAAATATTTGCATCTCTAGTGTCAATTATCAAAAGTCAACAAAGCC
AAAGCTACACAATAATTTTGCTATAACTCGACGCTTTAAACGCTCATGAAAATCAAAAGTAATTCTGATTAAATAAAAGTGAGAATGCCATAAACCATAGGTAAATTAGA
CTCAGAAATAAGGTCTGGTATGAGTTGCAATGCCTGTCTTGCCATGAGCATTTGACAAGTAATCAAGCACTTACGCAGAAGCTGTTGACTAATGGTAATGACACTTAATT
AGAAGCCAAGTAGTTCATTAAACACTTGCAGTTGACTTGCCACGCCATCGACGATTTGCCCCAGGATGCACCTACTTCCTGCCCCAAGTGTTTGTGGGCGGATCCTTCTC
TAACCAAACGGTTTATGTGTGCGTGAAATTGGTTTCCCGGCTTGGGCCAGGTCACCACAGTTGGCCTTTTGCTGCAGTTGCCACAGCTTTGGGTTTCGCAAGGTAATGCC
AGCTGATGTTTTCTTCGACAGCAGCATTGTTTGTGGTTTTTGGAAAACAGCTTAAGCGAGAAAATATGAGCGAAAAACTGGGAAGGAGCTTACGAAAGGTTCTGCAGAAG
CGATAGCCCAGAACGAGTGCGAAAAAATATGAAAATTAAATGGCGACCTTAAAGTGAAAGGGGTTCTTACATAAAATATTGTAGCACACTTGCCACTTTGTGCCACCCAA
GCAATAGCGTTTCCAACTAAAATGTTATCCCATCATATTGTAACCCGCCTCACCACATTTTTTCCATCCAGGTGTCGTACTTCTCCACCAGCCCCGAGCTGAGCAACAAA
CAGCGATTCGAGTACTTCTCCCGCACTATTCCCTCGGATCACTACCAGGTGAAGGCCATGGTGGAGATAGTGAAGCGGATGGGATGGAGCTACGTGTCGATTATCTACGA
GGAGAGCAATTACGGAATCAAGGTGAGTGTTTATGTATGGAACTTTGATTAAAATGGAAGAAAGGAGTGCATCTATGGCTGAGAAATGGAGTTAAAGAATGAGAAGCTCT
ATGTTTTAAAAGAGCTTGTTAAAATTCAAATTTAATGTATGTTGCTTCATTTTACTTTGGATACTTCTTGGATAAAATCTATTTTGGATAACCATTAAACATTGAAAAGT
CAATAATAAGTAAGTTCAATAAGTTTCAGCCACTTAAATTTCCTCTTTTTCAAAGGACTACATAAATTTCACCTAATAGGGCAAATCATTTAACTAGTAATCCTTAAACG
GTGGAACTTTAAAATAATGCCAAGTGCCGTTTAAAAGCGTTAGCAGCACTCCATTATCATATGTTTATATCCTTAGCCGCTTGCAGTGTAAACGATATACATTTTAAGTG
AACGCCTTGTATCCTATCCAACTTGTACCTGTGTAAACATTTTACCCATCCCACCACATGCGTTTTTATTGCTCACCCACACTGCACTTGTTTTCAGGCTTTCGAGGAAC
TGGAGGAGCTCCTCGCACGCCATAACATTTGTATTGCCATCAAGGAGAAGCTGGTGAAGGATTCGGGAGTGGCCGAGGACATCGCCTACGACAATATAGTGCAGAAGCTG
CTGACGAAGCCGCGGGCTCGAGGTATGTAACTGGAGAGTGGGAAGTTGGGATTGCAGAGAGGGATCGGAGGCGATGTGGCGGAGGAAAAGTGCATTTCCAACCGGAACAG
CGACTGCAGTCTGCCAGTCAAAATGCAATTTAAAATGAAAATGTCAAAGCGAAGGCGCAGTGCAAATGCGAAAGGAAAGCCCTTCGCAACGCCCGCACAATGAATGCGAT
CCGGAGTCGAGAACTATGCTAGGATTTTTGGCTGGGCGACGGAAATTGCGTAAATTAAATGAAAATATTTAAAAACTCCCTCGCCAGCGCGCCGTTGTGCCATCTGTGCC
AGCGGGAAATTAAAAAGTTCACGACAAAACAAACGGAAAAAGTTATTTGCAAATAAATTGAGAGATAGTCGCCCAGAAGTGGGCAAACTGATACGTCCAATACTGTAGTT
GTGCAAAAAAGTGTGTATTTGTTATGGCCAAAAATGAAGGTCCTGTCTATTGTCCGTGGACAAAAAGCCAGGAGTACTTGGCGTCCCTGAGGTCTGATGAATGACACCAC
CTATTCGACATGGCAAGCAATTAAAAATTCCAGTCAGGGTTGGCCAAGTTCAAATTTATCGGAACGATTTTTCATCCTAAATGAAAACGATTTTTCCACATGGCATTGCC
CCAATCTGCGCCAATCATGCGATTCAGCAGCGAGATAAACAAAAAATAAATGTCAATCAATTTGACTTTTCACATTTTGTGTTCCCCTGTCGAATATCAGCTGCCATTCG
CAGACGCCGAAAAGAAATATAAATTAAAAATAATTACCGCCCACAGCTTTGCGCATATCATTAGTCAGTGCCTCGAAAGGGTGAAGAATATCAGGGTCGTGCCCGCCAAT
AAAAAGCACACGGGTCGACCAATAACAATTGACTTTTGTCTGGCAAACAATTTTTGATTTACATAAATGCGCCATCAGGAGTCGTAAAGCAGTCGGTCGCCCAAGAAAAG
GGTTCCTCCGCCACCATTATTTATTGTGATTTGTGATTTGACGCCGCTTACAGGGCGCTAGCCGCCTCGCAACGATTAATCAATTTGACAAAGTATTTTCCCTGGCATCA
GTTGTTTAATTTTGCATATCGAGGGTTCTTTCACTGCTCCTGGTCCTGGTCGTGGTTCTGGTGGGGGCTCCAGTTGGGTCCCTGAAAAATGATGCCCTTCCGGATCCGTA
ACTGCGGCTGGTAACTGCCGCCAGGGGTGAACTTTACGCTTGTTCGGATCCCCACACATGTCACTCCGGAGCTAAATGGAATGTGGCGGGGGGCCGACATGTTTTTCCT
TTTTATCGCCAATGTGGCAAGCTGATGCGATTTCAGCCAGGACATGACCGGCCCCCACCCGCTTTAGATCACTTTCTGGCCCGGCTTTAAAGTTTCCCCCTAATCTGCAT
AAATATCCTGCCGAGTGGGCTCTTAAATTAAATTTTATATCCCTGCTAAGAATAACCGTGGTTCTCGACTGGTATAAACAGGCTACTGTGTACGAGAAAGCCGCTCAAGT
GCATGTCTAAGGCAAAGTGAAAATCCAAGGAAAATATCCTTAACTTCACAAAAGTATCGACAAGTTTTCCCGTTGTCGCCGCCGCTTCCTTTGAATAAGATTCCCTCCTC
TTACTTCGAGTGGCGGAGCGGCAGTGGCAAATCGATACCGGGACATGTTGAAGGGGCAAAAAATTATGGCTAAGACAAAGGACCTCTTGGGCGGACCGTTAAGAAGTGAC
AACTGCAGCTCACGTATCCCATACAAGTTTGTCCAGAGCGAAACAATAAGGGCTGAAAGCGTTGGAAAATTTGGTAAGCACCTAACCAAAGGATAGTTTTAGAAGCTCCT
GATTTCCATCCGTAGAACACAAGCGACTTAGGTGTCGGCTTTACCAATTGGCCATTGGGAACGTTGTTTGCTTGGAGTTTCTAATGATATCTTCACACTTTTTGTGGACT
TTCTGTTCGCTTCTTGGGACCAATTAAAATATTGTTTTCCTTGCAAATTGAAGAACAATACCTTGGCAGAAAATCTAAATATAGTAAAGATACGTCGCCAAATGGCGAAC
ATCTGAAAAAGTGGTTGCATTACATCAATATATACCCTTCGTCATGTGTACCCCTTGTACCCCCCATACCCTTGTTTTATAGACTTGTTTTACGCCTAACGGCTATAGTA
ATCCAACTTAATTTATATAGCATTTGCGTAATTAATATTCAAAGCGAAATAATGGGCTGTCATTTAATACTATAAAAACACTTTAATCCGCCCTCGGAAAGGAGAACACA
AGAGCAGTTGGCAGCTGCAGATTGTTTTCTTTTGCGCACAAAAGCCGCTCATTCACTATTGATTGCAATTAAAATTAAATTCTTTCTTCCTCCTTTCGCATCCTTTTAAT

ACGCCGCTGCCCAGGGGCCATTATCTTCGGTTCGGATCAAGAGGTGCGCCAAGTGATGCGGGCGGTGCGCCGGGCGAATGCAACTGGTTCCTTCTCCTGGATCGGATCCG
ATGGCTGGAGTGCCCGGAATTTGGTGTCCGACGACTACGAGCCGGAGGTAGGTGAACTATTCCGTCAGACCGGATGACGGGCCAGTTACCGTCAGCCGGAAATTATGAGC
CTAAATGGGCGGAACAACTCGGTTTGATTAATGACTCGGATTTCGGTCACATACTCACGCACGCACATACGAGGGCCCAACCCCATTACATGATAACCAAAAAATTATAA
ATGATGGCGGCCGAAAGAACACTTTCACACGCCCTTTCCCATGATGAATTATGGTTCAATTTGGATTAATCGTGCCTTTGTCTGTGCGTGGCTAATTAATTTGTCTAAGC
CCTTCAAGCCCAGAAATGGTCCTCCTCTCTGATCAAGCTCATTAAGTATACGCCGCGTGTGTCTTGGTGTCATCAACAGGTGGAAGGCACGCTGTCTGTGCAGCCGCAGG
CGAATCCGGTGCGCGGGTTCGAGGAGTACTTCCTCAGTCTGACGGTGGAGAACAATCAACGCAATCCCTGGTTTGTGGGTAAGTGAATTACCGATTAAAATTAAGCTGGT
GTTCATTAATGCTCTGATATAAGTAGTCCTTTATTAGTATCCCCTTTCGCTAAATGGTTATCTTGATTTCAGAATTCTGGGAGGATCACTTTCAGTGCCGATATCCCGGC
AGTACGAGCACTCCGTACAATAACTACACCAAACAGTGCACCACAAAGGAGCGACTTTCACGCCAGAACACTGACTTCGAGGACCAGTTGCAGTTTGTCAGCGATGCGGT
GATGGCATTTGCATACGCCCTAAGGTACGATTAAGCTTATTTTGGAAGCAACCCACTATCATTTTAAACGTATTAAGAAAATAATAGTTCTTTTGATCAACCCTTTTAGT
TTTTATCAACCTATGGAGTAGCAATAAATAAAATGTAATGTGATTTCCTCTCCTTAGGGATATGCACCGCGACCTGTGCGGCGGAGGTCCTTCGCTCTGCGAGGCCATGA
AGCCAACGAAGGGTGCAGACTTGCTGAAATATTTGCGAAAAGTGGAGTTTGAGGGTGAGTGAGTAATAGAAGCCGTGTATCCACTTTTGCCGGATAAAAGCAAGTTAAAT
TAACGGAGTAACTATTCCCGGGAATTTTCAAAACAGGATGTCCGACTGACTCAGATCCCTTTGCTCTCGCTTAGTTCCTTCAGTCGGCAAGTGATGCGCACATAGCTTAA
GTGGAAATTTCCTGCCAACCCGTCCTTGCAGTTGGGCCAAAAGCCCCGCACACACACTTACACACACAATGGCATATCTCAGCCTGCGAGAGTCCTACAAGGCGCATTGG
CCATGCTTTCGAGTGGAAAATTGACTTGCTTGTCATCGTCATTGAACTCGATTGAAATTCAATAAATTACACAGACAGACGGACGGAAGGAGCCAAAGGCCCTGACTGA
CTAGCGCAGATTGTCAATTTTCTGATAGGGCGAGCGGGCAAGCGGCAAAGTCGGCGAAGTCGTCTCTGGCAGGCGTGAAACCATAAACCCATTTAACTCCTATTCGATCA
ATCCCTCCGCAGGCCTCAGCGGCGACGAGTTCCGCTTCGACGGGAACGGCGACGGTCCCGCCCGGTACAACATCATCCACTTCAAGCAGTCCCAGGCGGGACAGTACCAC
TGGGTCAAGGTGGGCGAGTACACCGAGGGCGAGCTGCGGCTTAATATGACGGGTAAGATAAAGCCCCCAGTAGGAAATACTTAAAAGTATAATACCAGGAACCCTTCAAC
ATTAAGTTTACCTGACTGAATTCGTGCATCCTTAGTCTCCTGTAACCTAATGTATTCGTGCTCACTGCTCCTCATCAATACTAATTTTGATCTCATTATCCCGCTAACCC
AGAGGTGAAGTTTAAGCGGCTGAGTCCCAAGCCGCCGGAGTCTGTCTGCAGCCTGCCATGTCTTGTGGGTCAGGCCAAGAAGTACGTGGAAGGCGAGAGCTGCTGTTGGC
ACTGCTTCAATTGCACAACCTATCAAGTAAGTGAAGCCCATCTGAGGCCAGTAATTTATTTGTTGGCCCTTCGAATAGATTGAAATGTAGCGCCAAGTGAATTATTCCGG
AACTAACGCGTTGACGGGAGATGGGAAAAATGGGGCAAAAACAATTAAGCCATCAGGCCTGAAAGGGGCTGCAAGTGAATGGCAAACATGCCGCCTTTGTGCTCGCCCCA
TCCATTTATCAAGTGTTCCTTTGTGCGGTGCGTCCATTGCAGATCCGCCACCCGGATGACGAGACCCATTGCAAGCTGTGCAAGCTGGGCACTCTTCCCGACGCCCACAA
GCAGTACTGCCGCCCCATCCCGGAGATATATTTGCGTCCGGAGTCCGCTTGGGCCATCGGTGCTATGGCGTTCAGTGCGACCGGGATCTTAGTAACGCTCTTCGTAATGG
GCGTATTTGTCAGGTGAGCAGCGCATACACCGTGATAAAAACGGGTATAGGAAATTACTTTTAATAATATTTAATAGTTATTAATGGCTTTTATAAAACCATTTGAAACA
ATTCAGTCTTAAGATCGAACATTTAACAATGAAATTTGCTTTCTAAATCCAAAAACCCATCAGGATGTTATATTCCTTTTCTATTAAGACCATTTTTAATAATACATAGT
CCAATATTATCATAAAACCTTGTTAATGTATATATTTTCTCAACTTCAAATAATTTAGACAATATATAATACTTTAACAATAATTAAATAATTTGAAGATAAGGTAACGG
GTATATTATAGTCAATGCATTCGTTCTTGTGTTTATAAGTATAGTTATTGTTATACTAGTGACAAGCTAAAGGATCTAGATCAAGCTGATCTGAGTGCCAAATATTTCAT
GAGGGCCTGTTTTCTAGCTATGTTTTTCTCTGACTGTATTTATTGGCACACAATTGTCCGCTTGCATGGATATGGAAAGCAATTAAAAAGTTTCACCCTTTCCGGGCAAA
CTTCGCCACAAAGTTAACCGGCAGACACATAAAAATGATGCATTAATTGGCGCAAGAAAACGTCGTCGCAAGTTTGAAACTGAAGTAACGCCTTTCCCTCCTTCACCTTC
CACCCAACAGGCACAACGACACGCCCATAGTGCGTGCCTCTGGTCGGGAGCTGAGCTACATCCTGCTGGCGGGGATTTTCATGTGCTACGGTGTCACCTTTGCCCTGGTC
CTGAAGCCGACCAACATAGTGTGTGCCATTCAGCGGTGAGTGTGCAAATTGCCGAGATGATGCGTGTGAATTTTCCTCCGAAACTCCCTCTTTTCCGCTTTTCCGCTTCC
ACACAGGTTTGGTGTCGGCTTCTGCTTCACGGTCGTTTATGCAGCACTGTTGACCAAAACGAATCGCATTGCACGGATCTTCAAAGCGGGCAAACAATCCGCCAAGAGGC
CCTCCTTCATTAGTCCCAAGTCCCAGCTGGTAATCTGTGCGTGCCTGGTTAGTGTGCAGGTGAGTGTGTGGAGTGGTCTGGTGGCGCTCCTGGCGGTGACCGACGGGGCG
GATGTGGAATCGAGATGCCCTGCGATGACTCCGTGTCATTCTTTCCATTCACTCCATTCGCAGATACTCATCAACGGCGTTTGGATGGTAATTGCGCCATCACATGCCAT
GCACCATTATCCCACACGCGAGGATAATCTGCTCGTCTGCGACTCGTACATCGACGCCTCCTACATGATTGCGTTCTCCTACCCAATCTTCCTGATTGTTATTTGCACGG
TCTACGCGGTGCTCACGCGAAAGATTCCGGAGGCTTTCAACGAATCGAAGCATATTGGTGAGTAATTGAGCTTTTTATTTCCAGAAGAGCTATAATTCGGCGTAATTGGC
ATTGTATTGGAAGGGTCTGTACAAAAAGCCATGCATGGGTTTGCATTGTCTATAACCTGGTTCATTATAGCAGCAATGCTATTAACTGGTTTATTTAGTGCATAAAATAG
CACAACATGTATGAATGAACTCCTTATAATTTAGTTATGTGTTTTGTATTTGGGTAATTTGGTTATTTCGAAACAATACACGCATGAAATATTTATAAGTCTATATTTTA
TGCATTGCAACCTCAGAGTTCTAGTGCAGAAATGTTTTATAATTATGAAGCCATAATGCCCATAATTGGCCACTTTGATTGATTACTTAACCCTATTGTTACAGGATTCA
CCATGTACACGACTTGTGTGATTTGGCTGGCTTTTGTACCGCTCTACTTCGGAACCGCCAACCACGTTCCGCTGCGCATCACCAGCATGTCGGTGACCATCAGCTTATCC
GCCAGCGTGACCATCGCCTGTCTCTTTTCGCCCAAGGTGAGTGAGTCCGGATAGCTGGGAGTGGGTTGGGCTTCGGCTTGGGGGGATTCATATCCACACGGGTCCTCTGT
AAATTGATTGGTAATTGCGTTATGCATTAATTATTGCGCATTTATCTGGCGACCGTTGAATGATTTTTGGTGCCACAGTATGCAGAAGCCATCGGCGATTTTATTCCATT
TCATTCCGTTTATTCAGGACGCGTGGACGAGTGTTTCATTAAAATGCAATAAGACGGAAGATGAGGACAGGCGGAAGCCAAGGGTTTGGCAATTCAACAACCGGCTGCAT
CGATTGCGATGTGTTGCTTCCAACCATATTGTTTTTAATGCCTACCCTTCCTCTTGGTCCATATTCCTTGGTCCTTGCCCATTTCATGATGTTTCCCCGAGATGAAGCTT
AAATGTCAACAAGCCATGTTCTCCCACGGAGCTGAACTGGTCATGATCTCCTCCCAAAGGCGCGCGTCCTTTATGTTGGCATACTCGGTCCCCATGCCAACATATTTTTT
AATACGCCCTGCTATTAGCTATTAAGATGCAAGTAAACTGCCAGACAAGCCAAATATAGCCAGCCAGTTGAGTTCATGGACTAGATGATGCAGCATGGCAAAGAAGCTGT
TTACCCCGATATTAAACGTCAAGGAGGAAGCTTACATTTAAGCACTGATATCTTCATTACACGGAAACACGACCTTTGACAGCAATGTGCAATGTGAAACGCAGGAAAGG
AATTTTATAAGAAAATGTCGAATAGCTTGTTTGATTATTTTTTTTGTATAAGTATTCCAAGCATGAAAGTTAGATCAGAAAATCCAAGTTATATTGGCAAATAATGAAAG
TCTCTTAATTGTATGGCTATTAAAGGGGAAACACTTTTCTAGTTAGTTTCCAAAGGCTTCCATATAACACTTAAAGATGCCAGAAAGGACCTTCGCCAGTGTCTATGCCA
TTGAAGAGCAAGGAGGTAGAAATAGTAAATCAGCCAGAAACTTGATTAGCAACGGAGCGGAGCTCCCACAACCGCACACGTCATAGTTCTTGGTCCCGGAGCGGAGCAAG
CTACAGTTTATGGTCTGTCTGCATATTCCGGGTGCAAAGTAGAACAAAACCCGGGGAGGCGAAGGGCCGTAAACACTCAGCAAAAGGCTATTATTAAATGTCGGAGTGGA
GCATACTTCTTGGCCCAGTGGCTGCTAACCGGGCCAACTGTACGCTATCAGATGTCACCTAATGCACGGTTGGTTTGCGACGAGGAAGGAACACTGTACTGCGCACACGG
CAAATACAGTGTGCACAAGCACACGAACCGCACCGTACACCTTGAGCCATAAAAATGCAATTTACAACTGGTTGTGTCTGGCAAAAGGAAAACAATAAAAGGCTGCCATG
AATGCGGCATGCCACACAACAACTGGGGATCACCGGACCAAGAGGCAACAAACTAAGGCCCTGGCGCAAAGGTGCCGGCAGCCACGCCTCCTTCCGCAGGACAAAGGCTG
GCAACCGAGGTGGCAAGCCATTCCGACTTCCGACTGCCACGCGCCCTTTCTGCTGTCCGTCCGGCTCTGCCGCAGTTTGTTTTCGAACTTGCCGCGTTTGTTTGCGGTGA
AAACAATAACAAGGGGCGTCCAAATGGCAGCCAAACAATGTCAAAGCAAAGCGGAAATGACCCACCGTGCAGGCTACGTAACAGTACGCAGCGTATCGCAAGTTGAAACC
TGGAGCACTGCTCGCAATTTGCGCTAAATTTTCACTAAGCGTTTTATAGATGGTGCCATTCTAGTAAGGAGTATGTGTGCTCTGAACCAGGAACTCGCTCACGTGGCCAA
AAACAAATTCGATTTGGCTTTAAATATATATTAGAAAGTCAAGACACCTCAAGAAATTGGTGGGAAAGTGTTTTCATGTATTGAATAGTAATCTTATTAGCGGTATTTTT
TACACCTCACTAGTAGAGTGAAAGGGTAGACTAGTTTCGTTGAAAAGTATGTAACAGGTAAAAACAAGCATTTAAAGTATAAAGATTATTGATTGGAATCAATAGCAGAT
ACGATCTGACCATGTCCGTCCATTCGTATTAACGTCGAAACTAATAAGTTGAGATAAAGCATACTGATTCCAGAGACGTAGACGCAGCACAAGTTTCCGTATTCAATATC
CGTTCCGGCAGATCTTTCTATAGTCGGGGAACTCGACTATAGCGTTCCCTAAATTATATTTAAATTAATCAAACAAAAATGAACGAATGCGTGAACGTGTTATGATTTAA
TCCGTCTATTACATTGTTCATTGTTATATTATTCTTAAAGTGTAAGTATGGTAATATGGGGCTCAGGCGTGCACTTCTGCTCCCTGTAAGTGGCCCAAGTCGTTGGCCTT
TGCTTTGACCACATTACTGCTGGCCGTAGAAGCAGGAAATGGCTAATGACACTTGAATTTGGATTAGGAATGGGACTTGGTGGTGGGCGTGGCAGGTCACGCACTGACTT
CTAATCAAAGCCGTAAGATAAGCCATCTTGGGGAACAACGAAATTGGATATCGCCTCCAAATAGCTATTACACACAATGGGTCGTTCGATTAATGCCTCGAACGGGAAAT
TAACCTAATTGGAAAGACGCACTGCTTTTCTGACTCGACCCTACGCTTGTGTCCTTGCAGCTGTACATAATACTCATACGCCCCGAGCGCAACGTGCGCCAGAGTATGAT
GCCGCCTCGCTACGGCAACATGCACCGCACCGCTGGCACAGGACCCTCGTCCATGATGGCCGCCGCCGTGGTGACCGCCGCCACCTGTGCCCAGGAGGAGAAGATCCAGA
AGCATATCACTCCCACAAACACAGGTCAGTAATTAAGGGATATCGGTTAACGAGTCCCGGGTCAGTAAGCGTAGCGCTCAAAATACACACACTTCGTGTCGGGCAAATGT
ATTATTGACATATGTAATTTGCTGTTGAATAGTGCAATTAAGAGGTATAACAGTTGCTTGCCTCGTTAATCATTATCGATCAATTAATCAAATTATTGAAAGTGAGTCGA
TTTTATCTCCTGACAATGTATTAATCGTATTTTGCAATGAATTTATATCAATTAAATTATATTACTTTATTAACCCATTGCACCAGCTGTACCTATCTCTCGTATTCATT
TTTAATTTTTGTGGAAATGCTTCAATTAATTGAGAAAAGTTTATCATAAATTGGGTAATATAGCAAGTCTTAAAATAACTCAAGCTGTAAAAGCTGAACAGTTTTATCAA
AGTAATAGTTAGGCTGACAATCAAAAGGTATATCCTTCTTAATTTAATTGATTTTAAAATGACTAGGATAAAGAATAAAACTGAATTTTTTCCATTCACTACCCATAACT
TGCAGAAAACTCGCTAAAGACGAAGAAATTGTGCGAAATGGCCACTCAGACAATTTCCAGTATAATCACGTCGCTGGACATCAATGCGTACAATCAAATACCATATGCCG
ATCAATATGTGCCAAACAATGCCACAACGGCCACGACCACGCCCACAGATAAGGGCAGTGGCAACGGCAACAGCAACAGCAATAGCAATACGAATGGTAGTAGCAATAGC
AACAGCAACAGCAATGGCGATGGCGATGTGGCAACAGAGCAAGTCGTGGCCAACAATAATAAAATCAATCAGCGACAACATGGCCAGCCGGCCGTGGCATTTGCCATCAC
ATCGACTAACAATCACATAAGCGGCCCAGCAGCAACCATAACAACAACAGCGGCAACTGGCGAAACAACAACTGAGGCGCTGGTGGCTACAATAACAACTCCATTGGCAA
CGGTGGATGGCAGCGAAACCATGTCCGTGGCTCCCAACAATGGGCATGGAAACAACGGGAGCCAACGTCCGCCGGTCCTGCAAACAAACTTATAGCTAAACGCGCAACAA

FIGURE SHEET 3

TTGCAGCAACAGCTGCTGAAGGAGCAGCAGCGACAACAAGCCGAGCTGCAAACGCAACATCCGCAGCAGCAACATGGAACACAAAGCTCTTCCAGCGACACCGCCAGCAC
GGGCAGCGGCAGTGCGGCGAAGAAGCGGGCCGCAATTCCCGTTTAGCTGCCCCCGGAGATTCGAGATTCTCGCAACCATACGTAAGTAGAGGAAAGTGTCTGTTGACCCA
AGCGATTCCCGGCATCAGTCTAGGGATGCGGAAAAGCAGACAAGAGGAGGGCAGGAGGTCGTGCTAGGACGGGACTAAATTGGGAGCAGCTGGGCCACAGAAATATCTTC
AGATCTTAGAAGGCGAAATAATTCATTTACCACTTTTGTTCGCTTATTAATTTGAGTGAATTCTGGTTATACCTTACCCCCATAGACACAAATTATCTTAACGTTCTTTT
ATATAAGTTTATCGATAGCTGTACAACAATGTATCTCAATAATGAAGATATTCCTGTGGCACTTAAGGCTTTCTTCAGATTTTAACATGCACGCAAGAGTCATAAGTAAG
TAAGAGTCGGCTGATCAAAAACCTTCGCAACTCTTCCCAAAAACAAATGTGTTTCTCCATTT
(SEQ ID NO: 4)

Exon: 1001..1143
Exon: 3372..3542
Exon: 4058..4202
Exon: 6615..6757
Exon: 7120..7228
Exon: 7333..7504
Exon: 7648..7754
Exon: 8263..8412
Exon: 8583..8716
Exon: 8953..9143
Exon: 9801..9935
Exon: 10017..10179
Exon: 10294..10507
Exon: 10885..11036
Exon: 13701..13884
Exon: 14430..14572
Start ATG: 3468

Transcript No. : CT5032
GGTCCATCAGCAACATAGATGACGCGGAGTATCACTGCAATAAGACGCAGGTGCGCAAGGTCATCTCGGGGGTGGTGGGTGCCGCCTCCTCGGTCACATCCATACAAGTTGC
CAATTTATTGCGGTTATTCCGAATTCCCCAGGTGTCGTACTTCTCCACCAGCCCCGAGCTGAGCAACAAACAGCGATTCGAGTACTTCTCCCGCACTATTCCCTCGGATCAC
TACCAGGTGAAGGCCATGGTGGAGATAGTGAAGCGGATGGATGGAGCTACGTGTCGATTATCTACGAGGAGAGCAATTACGGAATCAAGGCTTTCGAGGAACTGGAGGAGC
TCCTCGCACGCCATAACATTTGTATTGCCATCAAGGAGAAGCTGGTGAAGGATTCGGGAGTGGCCGAGGACATCGCCTACGACAATATAGTGCAGAAGCTGCTGACGAAGCC
GCGGGCTCGAGGGGCCATTATCTTCGGTTCGGATCAAGAGGTGCGCCAAGTGATGCGGGCGGTGCGCCGGGCGAATGCAACTGGTTCCTTCTCCTGGATCGGATCCGATGGC
TGGAGTGCCCGGAATTTGGTGTCCGACGACTACGAGCCGGAGGTGGAAGGCACGCTGTCTGTGCAGCCGCAGGCGAATCCGGTGCGCGGGTTCGAGGAGTACTTCCTCAGTC
TGACGGTGGAGAACAATCAACGCAATCCCTGGTTTGTGGAATTCTGGGAGGATCACTTTCAGTGCCGATATCCCGGCAGTACGAGCACTCCGTACAATAACTACACCAAACA
GTGCACCACAAAGGAGCGACTTTCACGCCAGAACACTGACTTCGAGGACCAGTTGCAGTTTGTCAGCGATGCGGTGATGGCATTTGCATACGCCCTAAGGGATATGCACCGC
GACCTGTGCGGCGGAGGTCCTTCGCTCTGCGAGGCCATGAAGCCAACGAAGGGTGCAGACTTGCTGAAATATTTGCGAAAAGTGGAGTTTGAGGGCCTCAGCGGCGACGAGT
TCCGCTTCGACGGGAACGGCGACGGTCCCGCCCGGTACAACATCATCCACTTCAAGCAGTCCCAGGCGGGACAGTACCACTGGGTCAAGGTGGGCGAGTACACCGAGGGCGA
GCTGCGGCTTAATATGACGGAGGTGAAGTTTAAGCGGCTGAGTCCCAAGCCGCCGGAGTCTGTCTGCAGCCTGCCATGTCTTGTGGGTCAGGCCAAGAAGTACGTGGAAGGC
GAGAGCTGCTGTTGGCACTGCTTCAATTGCACAACCTATCAAATCCGCCACCCGGATGACGAGACCCATTGCAAGCTGTGCAAGCTGGGCACTCTTCCCGACGCCCACAAGC
AGTACTGCCGCCCCATCCCGGAGATATATTTGCGTCCGGAGTCCGCTTGGGCCATCGGTGCTATGGCGTTCAGTGCGACCGGGATCTTAGTAACGCTCTTCGTAATGGGCGT
ATTTGTCAGGCACAACGACACGCCCATAGTGCGTGCCTCTGGTCGGGAGCTGAGCTACATCCTGCTGGCGGGGATTTTCATGTGCTACGGTGTCACCTTTGCCCTGGTCCTG
AAGCCGACCAACATAGTGTGTGCCATTCAGCGGTTTGGTGTCGGCTTCTGCTTCACGGTCGTTTATGCAGCACTGTTGACCAAAACGAATCGCATTGCACGGATCTTCAAAG
CGGGCAAACAATCCGCCAAGAGGCCCTCCTTCATTAGTCCCAAGTCCCAGCTGGTAATCTGTGCGTGCCTGGTTAGTGTGCAGATACTCATCAACGGCGTTTGGATGGTAAT
TGCGCCATCACATGCCATGCACCATTATCCCACACGCGAGGATAATCTGCTCGTCTGCGACTCGTACATCGACGCCTCCTACATGATTGCGTTCTCCTACCCAATCTTCCTG
ATTGTTATTTGCACGGTCTACGCGGTGCTCACGCGAAAGATTCCGGAGGCTTTCAACGAATCGAAGCATATTGGATTCACCATGTACACGACTTGTGTGATTTGGCTGGCTT
TTGTACCGCTCTACTTCGGAACCGCCAACCACGTTCCGCTGCGCATCACCAGCATGTCGGTGACCATCAGCTTATCCGCCAGCGTGACCATCGCCTGTCTCTTTTCGCCCAA
GCTGTACATAATACTCATACGCCCCGAGCGCAACGTGCGCCAGAGTATGATGCCGCCTCGCTACGGCAACATGCACCGCACCGCTGGCACAGGACCCTCGTCCATGATGGCC
GCCGCCGTGGTGACCGCCGCCACCTGTGCCCAGGAGGAGAAGATCCAGAAGCATATCACTCCCACAAACACAGACGAAGAAATTGTGCGAAATGGCCACTCAGACAATTTCC
AGTATAATCACGTCGCTGGACATCAATGCGTACAATCAAATACCATATGCCGATCAATATGTGCCAAACAATGCCACAACGGCCACGACCACGCCCACAGATAA
(SEQ ID NO: 5)

Start ATG: 240

MVEIVKRMGWSYVSIIYEESNYGIKAFEELEELLARHNICIAIKEKLVKDSGVAEDIAYDNIVQKLLTKPRARGAIIFGSDQEVRQVMRAVRRANATGSFSWIGSDGWSARN
LVSDDYEPEVEGTLSVQPQANPVRGFEEYFLSLTVENNQRNPWFVEFWEDHFQCRYPGSTSTPYNNYTKQCTTKERLSRQNTDFEDQLQFVSDAVMAFAYALRDMHRDLCGG
GPSLCEAMKPTKGADLLKYLRKVEFEGLSGDEFRFDGNGDGPARYNIIHFKQSQAGQYHWVKVGEYTEGELRLNMTEVKFKRLSPKPPESVCSLPCLVGQAKKYVEGESCCW
HCFNCTTYQIRHPDDETHCKLCKLGTLPDAHKQYCRPIPEIYLRPESAWAIGAMAFSATGILVTLFVMGVFVRHNDTPIVRASGRELSYILLAGIFMCYGVTFALVLKPTNI
VCAIQRFGVGFCFTVVYAALLTKTNRIARIFKAGKQSAKRPSFISPKSQLVICACLVSVQILINGVWMVIAPSHAMHHYPTREDNLLVCDSYIDASYMIAFSYPIFLIVICT
VYAVLTRKIPEAFNESKHIGFTMYTTCVIWLAFVPLYFGTANHVPLRITSMSVTISLSASVTIACLFSPKLYIILIRPERNVRQSMMPPRYGNMHRTAGTGPSSMMAAAVVT
AATCAQEEKIQKHITPTNTDEEIVRNGHSDNFQYNHVAGHQCVQSNTICRSICAKQCHNGHDHAHR*
(SEQ ID NO: 6)

Name: metabotropic glutamate receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384693
CAGTTAGTTCCTAGTTAACAAGTGAGGTTTGGTTGAAATCTTTCTTTGCTATATAGGTATTTCATTTCCGTATGCTGTATCGGTTTTGTAAAATAATAAAGCATTTCGCGAG
CTCTCTCTCTCTCTGTATCTCAAAAATATAAAGTAAATGCCCCTATGTGTGTGCTTAAAGCATCTAGATTGATGAAAATGTACACTTCAAATGCTTCGCTCTGGTTTGCTTA
CAACGTAGCTGAAATGCAGCAACAAAATGGCAATGTATTATTTCCCTTTACTTCTTATATTCCTCTTGATTCGCAGTTGGAAGGAATTTTAGTTGTATAGGTGGAAAACGGC
AGCTGGATAGCAAGAAATGACAATCTTTCAAAACGTAGGATATCTTTTTGGTTTATGAAACACAAAACAAATCTTAAGCCTGTGTGTATATATATATATATCTTACGAGCAT
AATCAGCGTAAATCAAAAACAATTGATTTCTTACAAAATACATAAAACATAATTCGCCTGAATTCTATATATCATGTACATACAGTTCCTAGATCTTATACATATACCTACA
TCCATATATATATACTTGATTTTCTGGTTGTTTAAATTAGTTTTTAAGTATTTAAAAATGGCATGTGGTGACATGTATTACAATACATCATACATATATGGTACATACATTT
ATATTTAATATCGATTTCCAAGTTGGTATAAAATGCATGCAAGTCGAAGGGCCCGATTCCAATTCCAATTCCAACCCCCCATCGCTCGAGTTTTGGTTAAGCTTGCACGGCA
CCGCAGGGGTATGGATGTGCCCTATGGTCTAGGTTTTAGTTATTTCATATACATTTATAAAGTTTTGGGATTTTCGAATCACTGCACGTATAGCATGACAGCATGTTGTTTC
ACACAATGCTCCATATTTCGACTAACGACTAATACTAGGTATAGCTAGAGATCCTAAGCTACAATTTTTGCCTAAGACTGTTCCGGGGGAGTTGCGTTAACTGATCACTTAG

CCAGTGGTTCCAGATAACATCGCACCGTGTTGTTCTCATCGAGCATGGCCGACAGGTGGCGCTGCTGATGGTGCGGAGGATGCTGATGGGGATGCGGATGGGTTAGCTGGTG
GGGATGCGGATTGGGATTGCTGTGTGATGTATGCGACGACGATGATGTGGTGGACTGTGGTGACAGTTTCTGTAGCTGCTGGCGCTTCTGCGCGGGCGTCTGATTCGCCGGC
GGCAGCGGTGGTCCGAGTTGCCGTGACCGAATAGCTCATTTACCGCCAAACGGCCGCCTGGCGACTTGGCATTCAGTGCTTGCTGCTGCTGCTGCCGGAATGCGCCCACGCA
GGGAGCCCTCACCAACGCCGCCGCTAACCCCAAAAACCGCACCGTTGTGCTCGGGTTGCGGTCCTGCTCCTGGTGGCAAAGCCCCGTAGTCAAGAGGCGCTCCATGTTTGCG
CAGCTGACGTCTGTAAAGGGGATAAATGTAATTTTTCGTAACGTTGATATAGTCAAACATTAGAAGTACTTACTGAAACCGCCGCTTAAAATCCTCGTCCTGCATCGCCAT
CAGACGACGCATCCGCTCATCGCTCTGCGGTGGCAATCCATGAGCATCATCCTCCAAGTCGCTCTCGTCCTCCTCCTCATCCTCATCACAGACGCTGTCGTGCAATTGCTCC
GCGGCCGTTGCCTCCTCATCCTCCTCCTCGTCTTCGTCGTCATCGTCATCCGCCGTGCAGCGTCGAGCCGATCTTTGCTGAGCCAGCGACTGTTGTTGCTGCTGCTTCTTGA
GCTGCTGCTGTTGGCTCCTGGACGAGACACTGCCGCCACTCAGGCTTCCGCCGCCGTACCGACTACCGCCGGCGGCCACCGCCGCTCTCATGTTGTTGAAGACGCCAGTGCC
AACGCTATTGTTGTACACCTCGTCCACATCGAAGTCACCCACCTCGTAATGCTGCACATCCGAGTTGCGGTTGCTGAGGATCTCCTCGTAGACCGGCCCACCATCCGCGCCG
CCGACGCTGCCTTGGTTGTAGTAGGCATCGCCCATGCCCAGATCACAGTTGGAATGCTTATTTCTATACTGGCGCTGCATAAACTCCTGCTGGTACTGCGCCATGTGTGGCG
GTGGAGCAGGGAGCTGCTGCTGGCTACCAGCTCGTCGTTGCAGGTGCAACAGGTGAGCATGTTGCTCGGCCTGCTCCTCGTCCGAGGAGAGTTGCTGCTGCTGCTGCTGTTG
CTGCTGCTGATGTAGCTGCTGGCGACGCAACTGTTGCTGCTGCTGGTAGTAGGCAGCCGCATTGCCAGTGGCTAGTGGTGGGTCCAACGCCTCCGCGTATTGCGGATAGCCT
GTTGGCTGGCCAGGCTTGGTCTTCTTGTACGATGGATAGAAAAGCACCTGGGAGGCGGCGGCCGTGCGACTGTGCAGAGATCCACTAGGCGAGGCGGGAACTCCTCCGGCGC
CCCTCTTCTTTTTGCCATTGTGGCCGCTGTTCGGCTTTTGAGTCCAGTAGAATACCTCGTGCGCCTGTTGCGGCGATGATGCAGCAGCACCCGCATTATTGTAGTACTCCCG
CGGTGTTTTCAATCCTCCTGTTCCAGTTGATGGTTGCTTCCAGCTGTAGTAGTCCGGCTGATGGTAGTAAGACTCGCCCAAAGGTCCCTGTTGCTGCTGTTGTTGTTGTTGC
TTGTGCGCAAAAATCAGGTGTGTGGACCCGGCAGAGCTGGTGGGTGAGGCATGACCAGTGGAGCGGAGAGTGCTACCTCTTTCGAGCACAACATGGCCACCAAGGGTCTGCG
GTGGAGCCTGTAGTGTGCCCGTCATTAGGCCACTTTTAGGGTTCTTCTTGCTGGGCCTCGTTCAGTTCACAATCAGAGGTGGCTCCCATAATGGCATCCTCTGTGATGGG
CACTGGTGTCTGCTGTTGCTGCTGGGGATCGTCATGCGCCTCTTCGCTCACTCCTAACGGCAGCTTGGGCTTCTTGGACGTGGAGACGCTGCAAAGGGTTCCGGCTGTGGGT
CCATTGCCGGTAACAATGCCCGAGGAAATTGATGTTTTCGAGCAGCGCAAGCACTTGGGCAGCCAAGCGTGCTGTCTCACATACTTCCGATACTCCCGCCGGATCTTTTCGT
TTTGAATGCAATGGAACACGAAGATATAGAGCCCCAGCAATGTGTTGAAGCAGATGAAGCAGTATCCATACACATCAGCCGTGTCATCGTCCATTTTGGCTCCTCGCAAATA
GAAGTAGGCCGAGCACCAAACAGCGCTGAGCAGCAAGAGGAACACAAAGGAGCAGCGTATGTCGAACCTAGGATTTGAGGTACATGGGCATCGAAGGATCTTTAAAGTGTAT
ATGGATTTCAATATCTTACCGCACGCTAGCGAGGCGAGTATGTTCCTTGGTCTTTAGACCCGTGCGACTTTTGCGGCACATTATAATCCAGGAGAGGAATGTGTAACCAAT
GGCAGCCTGTAAAAGTGAATTAATTTAAAATGGTAACTGCGCTTTTGAAGCTTCATCTCATATTATGTAACTTACCACAAAGAAGACAAGCACTGGTATTACAAAGGTGGCA
TAAAACAAGGCATTCGCCTCCATCAGCACGCAATAATCGTTTTGGGTATAGGTGCTGGGATCGATGACCAGCGAGATGGCCACCACGCTCAGCGACAGTCCGTAGGACAAGA
GGTAGTAGCAGTTCACCTTGGGCGTCTGGTCCACCTCCAGCAGGAGCTCGTCCGAGGTGAGCGTTGAGTACGAATGGAAGGCTGCAGAGGCAAAGGGTTAGGGCTGCAAAGT
GTAGATGAAATTCGGCCTTCTTACCTTCGTAACAGAACCAGGCGGTGCCCGATAGGATGGCACAGTGTAGGAAAATAGTAATGAAGCCGCAGAAAATGCTTGTTTCGGTCTG
TTCAATGCCCAGGAGAAAGAGCAGCTCGATGGCCAGGAGGCAAAGGTAAATGCTGGTATAGATCGAGGTGCGCGCGGACTGGAGTGGACAGTTAAAGAGAAGACAGAGTGAA
AAGTTAATATAAGTTCGAGAAGTCAGTGCAAATTTTTCTTTAAGCAGCACCAGCAGAAGGACCCTAAAGATTCGATCATCAGGGAGCTTTTAGACATTTTAAGGAAATTGA
TAATCTTAAACGTTTTGAAGAATAGGTGTCAGATTTGATAGTCAAACTTAATTTAGAACTTGATTTACAATAATTTTGGTGGGTGGTGGTGAAAATCCAAAGACTAAGAGTA
ATACGAATACGAATAAAACAATGAACAGAACACGAAACAAGCACAACACATAGGTAATATACATACAAACTAAATGAACAAACACAAAATGCAATATTTACAGGACAAAATT
AGCAACATTTGTATATTTTATTAATTAATGAAATACAAAAGAAAGTCAGAAGTGTTCTAGGTATTTAGTTGTACGATGTCTATGTTAAATTTGTTAAACAATCAATGCACAT
AAAAGGGGTTATGGTATTTCAAAAATGATTAAACGATTAAGTAAAACATGCCAGAGTGATGATTTGTAACAAAAAGAAAAATGAGTCTATCATATGAAGATATTTCATTATT
GGTTTTGTGTGCTAATCTTAAGCTATGACAAATTACTACGATACATAAATATAAATGGGAGTCTTCGCGTGCAAGCTGTTACTGGGACGCAACAGTGATCATCGCTGCCCAG
CACAGCTTTTGAGTCTTATCTTTGAATCCTGCCAGAGAGAGGTTCATTAACAATTTAAGTTACTTAAGCATTATATCACATGTGTCTCTCGGTCATAAATACATATGTCTAG
AAAGCCTAGCCTAGCCTAAAATTGTGTTTATCTGTTTTTCACCACCACCACCACTTATATTGTAAGTCGTCCATGTTGGCCTGCATAATGATGTTGTTCGTGTTATCCAAT
CCATTGCCGTCCACATTAAGGTTGTGATTGCGGTGCGGATGTGGCTGCGGCTGCTGGCCACGCTGCTGCTGCTGCTGCTGGACATGCTGCTGCTGCAGATAGTTGTGGC
TGATGTTGAGGTTATTGCTGGCGTTGCTGGCCTGAACTGCATGTTGCTGCAGTGCAGCGGCTATTGAGGCAGCAGCAGCTGCAGCAGCAGCCGTCTGCTGACCGTGCTGATG
CAGATTCAGATTCAAGTTGTTCATGGCATGTTGCATATTTCGCTGCGGCTGATTGGACAGCACAACAGCAGCAGCAGCTGCCGCAACAACTTGTTGTTGCTGTTGCTGCTGC
TGCTGTTGGACATTGTAGTTCAGATTGTTGCGGTGTGAGTTGCGGATGGAATTGTGTTGGATAAAGTTCGTGTTCCCATGACTGGGCGGCGGCACATTGGACTGACTGGTTG
GCGTGGTCAGGGTCAAGGACTCGTGGGTCTGGTCGCGAATATTGTTTTGCTGGCGTCTGCTGCCCGACCGCTGATGCGGCAAGGGATTGGAGCCGTTTCTTACCTTCACAAA
GACCCCATTGAACAGCTTCAGCGTTAGCAGGGCGATAACTATGAAGACCACGCAGATGCCGATGCTTATGTAGATGAATATGCGCATGTTTCCATCGAACATGGTGAACAAC
GAATGCTGGTGCTCATCCACAACGTCCATTAGTATGGCAAAGTTTGTCAGGTGGTTGCAACTGCAGACGCTGTGCGTGCGGTTAGTGGACTCCAGACTGCATCCGTTGGCAG
ACCACGCACTAAAACAAAAACATTAATTCGTATAAATTATTAGGCATATGAGCATATGAGTCTTACTGGTCAATATAGTTCCAGAACACGCAGGTGGGATTCGTTACATTCT
CGGTCTTCAGATGTTTCAGTGTCAGGGTTATGGGCTGGGAGAGTTGTATGTGACGCCCCTTGCCCAAGCTGGCCGAGATCACCTTGCTGTTCAGGATGCGTAGGCGCTGTTG
GATCTCCCCCGCGTTGACATCGCTGTCGTTGCTCAGGATGGCTGTATTGCGAACTAAGAAAACATTTAGATTATACTATTCTAACGTTTTTCAGTTCTCTGTGGACTTACCG
TAACTGCGGGAGCTCTTGAGATCGAAGTGATCATAGCTGGGCTTTAGAATGGATTCCAGGCGATCGAAGGCGGCGAATACAATTCGCACCAGACCGCCTTCACTATTATCTA
TTAGAGCAGCTCGTGGCAGCTCAATACGATCCGAACTCAAGGGCCACTGATCCGAATCTGGGAAGACCACGCTGGACTGGATAGTCTTGGTTTCCAGAACTCGAACGGAGAG
CACTATGGTGATATACAGTTATAAGTATTTTGAAATGCAAACAATGCAAACATTTAGAACTCACATATATTTTTGACTTTTTGCACCACGCTGCGCTCCCTGATGATCGTAT
CCGCCAGCAGAAAGGCATTGTATTCCAGGCCAGTTAGCAAGGATGTGGCTACACGCATTTGGTCCTCCGGATTGAGATCCAACCACGAGGACAGCTGCGATTCGTCCAGCAG
GTTGGAGCCGGTTTTGACCACACAATGCAACAACTCCATGATCATAGCCTCGCGCTGTCGCTGATCCGGGAAGGTCTCCTTGTCGTGCATCATCTTCTCGGACACTGTTTGG
ATAATCTTCGTGGTGACCAACATGTCGCCGCCGTAGAGCGTTTTGCTACTGGTCACTTCGGACATATCGTTGGCGATGGAGATCAAGGAGGAGTCCCGCTGGTTTACTCGCA
TTTCCAGATTGTTGAGCCAAAGGCTGCGGCATTGCGTCAGATCAGGTGTCGCGGGATGCCAGGTGGGCTCAAAGTTGCGCAGACGCTGGTTTACCTTGTGGGCCATGCTCAC
CGGCGGCTCGCAGCTGCTGCTGTTGTGCAGACAGTCGCCACCGCTGGGTGCCGGAGTTGTCGAACTGATGTCATCATCGTACTCATCGTAGCCGGAGTCGGGTATCCTCTTC
ATTAGAACGCAACGCCACTTGGCGATGCCGGCTGCTCCGCCAGGACAGGGCTGAACATTCACATCACCCACTCGAGTCATGTTCCAGTACAAATTGCGGGCATGAGTGGGGC
CGCAAAACAACTTGGACTCATCATCGCCTCCGGACATCCCGGCTGCGTTCAAATTAATATTGTTGATGGTGCGCACCACGCTGCCGTTCCCGGGGACAACGGATCCACTGGC
CGCAGTGCCCGGTGGGCGGTTGGGTGAGCTCTTTGTGGTCAGTAGGGTTCCGTCATCAGTTCCGGAACCTCCGACGCCCGTGAGTATCCTCGTGTTGGATGGTGAAGTGGCA
TTACCGCCAGCCGGCAGCTTGCTGCTCGGCTTGGTGGGTGAGGCAGTGTCTTCACCGTGGTGGTTGTGGTGGTGGTGCAGTTGCGGCGGCGGTGGCAGACCATCATGGCGGC
CAGCCGGATGCTTGGTGGTCGTCGTGGAGGTGGCCCCACCCTTCAAACGTCCACCGCCCACGGCTGTCGTGCTGCTCGAGGGCGTGGAGTGCGTGACGGCGGTGTGTTGCGG
TGGTACGTTGAACAGGCCGGGATTCCCACTGATTCCCACCACTCCGGGAAGTGTGGAAGTCTCGGCGGGGGCGGCGCACCCGCTCCAACCCCGGGCGGATGGATCAGTCCA
CTGCCGTTGCCAAAGATCGGCGGACCATTGCTCAGCACCCATGGCGGCGGACTGGGCCTGTTGGTCGTCGTCGAAGTTTGGGCTGCACTTATGCACTGGTAGTGTGCCTCCA
GATACTTGTGGGTACCGGGACAGGGATCCCCGAACATGCTCGTGGCTGCCAACACGCCGCAGCTCTGCTTGTGGGCACATCTGGAAAATGGGAAAAATACGCGTTCATCAGT
AAACAGAGTTGGGTAATTTTCACATATTTAAGATGCTTGATTGATTGACTTGTGTTGAAATAGTTTATTGATATTCGCAACCAGAGTATGAAATTTCCAAATGACATGTTTA
ATGATTTGAAGATCCAATAAACTTTAAAAGATGTATTACCTTGAGTTCAGTACGCTGAGTGACTTGGGAAACATGCAGTTAACACTCCACTCCACATTCCCGTGGTCATTGC
AGATGGTAATCGAGAAGCGGCCATAGTTGGCCCGAATGAGGTTGATCACATCGCCGGGATCGCACTCGATGGTCAGTTTCTTACCTTCGCAGGCGTAGGCGGTTTGGTACTT
GGACACTGAAAGGCGAAGGCTTCATAAGTCACAGGGTGCTGGGCAGGGTATGGGTTTATATGGACATGGTATATATATATGGCATATAAAAAGAAATTTTTTGGTTAGCGGC
GATGATTGTTGTAGTTTGGTTTGGTTTGGATTGGATTGTTGGTGGGCCTGCAAAGACACTTACGATCTAACGATATATGCTCATAGGATATTGACAATATGGTTGGTAGCAT
GATACGCTTTTACACTGCCTGGGATATACTTAGTAATGCTGAGGTTGATGGTGGAGTATGTGGGTCGTTAGCTTGGTAGCAAGTGCAATAGTAGTCATAAATACCAAATCCT
GCCGGCCCAGATCACGAATCATCGATCGTCCATCATTGGTTCTTATCGGCAATGCTCATTCGGCAGTGTTTCACTCGGTTTGAAGTATTCAAACTGAAAGGGAAGATAACAT
TTATAATGTTTGGTATTATTATTATTATTATTTTAAGGCACATAACTAATAACTTAAGACATGTTGTATTTCGTTAGAAAGGAAAAGGAGAACTATCCTTTTCTTTTATTGA
TATTGGTTCCATTTAATTTACTTGTGCAATATCACATCAATGTTGAGTCAATCTAATGACCTGGTTACAGTAAATATCAAAAGTGTAGATTCTAATAAATTTAAGTTATTTT
AATATTCAGAACCTCTTTGTATAACTCCTATTAAAACACTATATAAGCAATGCAAATATCAAGAAAACATAATTAATTGTAAGAATTATAAATTGCATTTATATATCTTTAA
TAATTAAAATGGTCTTATATTTATGTATGATGATATCCAGGCCAAGCACACCTATGTACATATGTACATACCTCTTATGTATTTCCGTTTGGAGTCATATTCATGAAAGTTG
GACTCAAAAGTACCATTCAACTTATGTATCTGAAAGGTAGAGAATCCATTAGATAATTGTTTACATATTGACTCCTTGGGCAAACACGAACAAGCAGGCAAAAGACAAATAG
AAGATAAGAAGATGGCATGGGGCAGGGGAGGGGGTAGGGTAGGGAGCTAAGAATACACACCCTTCTGTGTACACCCACACTGTGGTTGTAGTTGTTGCATGGCCATCGATTC
GTGGACAGAACAAAGACCCGTTTTCTTTGTACACCCGATGCACACTATTCACGAACATATATGCATGTGGATGTAAATGTAGAGAAAAGTTGCTGCAGCTGCGC
(SEQ ID NO: 7)

FIGURE SHEET 5

Exon: 8400..8352
Exon: 8181..7992
Exon: 7808..6337
Exon: 6274..6047
Exon: 5977..5779
Exon: 5720..4860
Exon: 3972..3833
Exon: 3777..3548
Exon: 3478..3381
Exon: 3315..1531
Exon: 1466..1001
Start ATG: 8400 (Reverse strand: CAT)

Transcript No. : CT8755
ATGCTACCAACCATATTGTCAATATCCTATGAGCATATATCGTTAGATCTGTCCAAGTACCAAACCGCCTACGCCTGCGAAGGTAAGAAACTGACCATCGAGTGCGATCCCG
GCGATGTGATCAACCTCATTCGGGCCAACTATGGCCGCTTCTCGATTACCATCTGCAATGACCACGGGAATGTGGAGTGGAGTGTTAACTGCATGTTTCCCAAGTCACTCAG
CGTACTGAACTCAAGATGTGCCCACAAGCAGAGCTGCGGCGTGTTGGCAGCCACGAGCATGTTCGGGGATCCCTGTCCCGGTACCCACAAGTATCTGGAGGCACACTACCAG
TGCATAAGTGCAGCCCAAACTTCGACGACGACCAACAGGCCCAGTCCGCCGCCATGGGTGCTGAGCAATGGTCCGCCGATCTTTGGCAACGGCAGTGGACTGATCCATCCGC
CCGGGGTTGGAGCGGGTGCGCCGCCCCCGCCGAGACTTCCCACACTTCCCGGAGTGGTGGGAATCAGTGGGAATCCCGGCCTGTTCAACGTACCACCGCAACACACCGCCGT
CACGCACTCCACGCCCTCGAGCAGCACGACAGCCGTGGGCGGTGGACGTTTGAAGGGTGGGGCCACCTCCACGACGACCACCAAGCATCCGGCTGGCCGCCATGATGGTCTG
CCACCGCCGCCGCAACTGCACCACCACCACAACCACCACGGTGAAGACACTGCCTCACCCACCAAGCCGAGCAGCAAGCTGCCGGCTGGCGGTAATGCCACTTCACCATCCA
ACACGAGGATACTCACGGGCGTCGGAGGTTCCGGAACTGATGACGGAACCCTACTGACCACAAAGAGCTCACCCAACCGCCCACCGGGCACTGCGGCCAGTGGATCCGTTGT
CCCCGGGAACGGCAGCGTGGTGCGCACCATCAACAATATTAATTTGAACGCAGCCGGGATGTCCGGAGGCGATGATGAGTCCAAGTTGTTTTGCGGCCCCACTCATGCCCGC
AATTTGTACTGGAACATGACTCGAGTGGGTGATGTGAATGTTCAGCCCTGTCCTGGCGGAGCAGCCGGCATCGCCAAGTGGCGTTGCGTTCTAATGAAGAGGATACCCGACT
CCGGCTACGATGAGTACGATGATGACATCAGTTCGACAACTCCGGCACCCAGCGGTGGCGACTGTCTGCACAACAGCAGCAGCTGCGAGCCGCCGGTGAGCATGGCCCACAA
GGTAAACCAGCGTCTGCGCAACTTTGAGCCCACCTGGCATCCCGCGACACCTGATCTGACGCAATGCCGCAGCCTTTGGCTCAACAATCTGGAAATGCGAGTAAACCAGCGG
GACTCCTCCTTGATCTCCATCGCCAACGATATGTCCGAAGTGACCAGTAGCAAAACGCTCTACGGCGGCGACATGTTGGTCACCACGAAGATTATCCAAACAGTGTCCGAGA
AGATGATGCACGACAAGGAGACCTTCCCGGATCAGCGACAGCGCGAGGCTATGATCATGGAGTTGTTGCATTGTGTGGTCAAAACCGGCTCCAACCTGCTGGACGAATCGCA
GCTGTCCTCGTGGTTGGATCTCAATCCGGAGGACCAAATGCGTGTAGCCACATCCTTGCTAACTGGCCTGGAATACAATGCCTTTCTGCTGGCGGATACGATCATCAGGGAG
CGCAGCGTGGTGCAAAAAGTCAAAAATATATTGCTCTCCGTTCGAGTTCTGGAAACCAAGACTATCCAGTCCAGCGTGGTCTTCCCAGATTCGGATCAGTGGCCCTTGAGTT
CGGATCGTATTGAGCTGCCACGAGCTGCTCTAATAGATAATAGTGAAGGCGGTCTGGTGCGAATTGTATTCGCCGCCTTCGATCGCCTGGAATCCATTCTAAAGCCCAGCTA
TGATCACTTCGATCTCAAGAGCTCCCGCAGTTACGCCATCCTGAGCAACGACAGCGATGTCAACGCGGGGGAGATCCAACAGCGCCTACGCATCCTGAACAGCAAGGTGATC
TCGGCCAGCTTGGGCAAGGGGCGTCACATACAACTCTCCCAGCCCATAACCCTGACACTGAAACATCTGAAGACCGAGAATGTAACGAATCCCACCTGCGTGTTCTGGAACT
ATATTGACCATGCGTGGTCTGCCAACGGATGCAGTCTGGAGTCCACTAACCGCACGCACAGCGTCTGCAGTTGCAACCACCTGACAAACTTTGCCATACTAATGGACGTTGT
GGATGAGCACCAGCATTCGTTGTTCACCATGTTCGATGGAAACATGCGCATATTCATCTACATAAGCATCGGCATCTGCGTGGTCTTCATAGTTATCGCCCTGCTAACGCTG
AAGCTGTTCAATGGGGTCTTTGTGAAGGTAAGAAACGGCTCCAATCCCTTGCCGCATCAGCGGTCGGGCAGCAGACGCCAGCAAAACAATATTCGCGACCAGACCCACGAGT
CCTTGACCCTGACCACGCCAACCAGTCAGTCCAATGTGCCGCCGCCCAGTCATGGGAACACGAACTTTATCCAACACAATTCCATCCGCAACTCACACCGCAACAATCTGAA
CTACAATGTCCAACAGCAGCAGCAGCAACAGCAACAACAAGTTGTTGCGGCAGCTGCTGCTGCTGTTGTGCTGTCCAATCAGCCGCAGCGAAATATGCAACATGCCATGAAC
AACTTGAATCTGAATCTGCATCAGCACGGTCAGCAGACGGCTGCTGCTGCAGCTGCTGCTGCCTCAATAGCCGCTGCACTGCAGCAACATGCAGTTCAGGCCAGCAACGCCA
GCAATAACCTCAACATCAGCCACAACTATCTGCAGCAGCAGCATGTCCAGCAGCAGCAGCAGCAGCAGCGTGGCCAGCAGCCGCAGCCACATCCGCACCGCAATCACAACCT
TAATGTGGACGGCAATGGATTGGATAACACGAACAACATCATTATGCAGGCCAACATGGACGACTTACAATATAAGTGGTGGTGGTGCATTTACCTTTGCCTCCTGGCCATC
GAGCTGCTCTTTCTCCTGGGCATTGAACAGACCGAAACAAGCATTTTCTGCGGCTTCATTACTATTTTCCTACACTGTGCCATCCTATCGGGCACCGCCTGGTTCTGTTACG
AAGCCTTCCATTCGTACTCAACGCTCACCTCGGACGAGCTCCTGCTGGAGGTGGACCAGACGCCCAAGGTGAACTGCTACTACCTCTTGTCCTACGGACTGTCGCTGAGCGT
GGTGGCCATCTCGCTGGTCATCGATCCCAGCACCTATACCCAAAACGATTATTGCGTGCTGATGGAGGCGAATGCCTTGTTTTATGCCACCTTTGTAATACCAGTGCTTGTC
TTCTTTGTGGCTGCCATTGGTTACACATTCCTCTCCTGGATTATAATGTGCCGCAAAAGTCGCACGGGTCTAAAGACCAAGGAACATACTCGCCTCGCTAGCGTGCGGTTCG
ACATACGCTGCTCCTTTGTGTTCCTCTTGCTGCTCAGCGCTGTTTGGTGCTCGGCCTACTTCTATTTGCGAGGAGCCAAAATGGACGATGACACGGCTGATGTGTATGGATA
CTGCTTCATCTGCTTCAACACATTGCTGGGGCTCTATATCTTCGTGTTCCATTGCATTCAAAACGAAAAGATCCGGCGGGAGTATCGGAAGTATGTGAGACAGCACGCTTGG
CTGCCCAAGTGCTTGCGCTGCTCGAAAACATCAATTTCCTCGGGCATTGTTACCGGCAATGGACCCACAGCCGGAACCCTTTGCAGCGTCTCCACGTCCAAGAAGCCCAAGC
TGCCGTTAGGAGTGAGCGAAGAGGCGCATGACGATCCCCAGCAGCAACAGCAGACACCAGTGCCCATCACAGAGGATGCCATTATGGGAGCCACCTCTGATTGTGAACTGAA
CGAGGCCCAGCAAAGAAGAACCCTAAAAAGTGGCCTAATGACGGGCACACTACAGGCTCCACCGCAGACCCTTGGTGGCCATGTTGTGCTCGAAAGAGGTAGCACTCTCCGC
TCCACTGGTCATGCCTCACCCACCAGCTCTGCCGGGTCCACACACCTGATTTTTGCGCACAAGCAACAACAACAACAGCAGCAACAGGGACCTTTGGGCGAGTCTTACTACC
ATCAGCCGGACTACTACAGCTGGAAGCAACCATCAACTGGAACAGGAGGATTGAAAACACCGCGGGAGTACTACAATAATGCGGGTGCTGCTGCATCATCGCCGCAACAGGC
GCACGAGGTATTCTACTGGACTCAAAAGCCGAACAGCGGCCACAATGGCAAAAAGAAGAGGGGCGCCGGAGGAGTTCCCGCCTCGCCTAGTGGATCTCTGCACAGTCGCACG
GCCGCCGCCTCCCAGGTGCTTTTCTATCCATCGTACAAGAAGACCAAGCCTGGCCAGCCAACAGGCTATCCGCAATACGCGGAGGCGTTGGACCCACCACTAGCCACTGGCA
ATGCGGCTGCCTACTACCAGCAGCAGCAACAGTTGCGTCGCCAGCAGCTACATCAGCAGCAGCAACAGCAGCAGCAGCAGCAACTCTCCTCGGACGAGGAGCAGGCCGAGCA
ACATGCTCACCTGTTGCACCTGCAACGACGAGCTGGTAGCCAGCAGCAGCTCCCTGCTCCACCGCCACACATGGCGCAGTACCAGCAGGAGTTTATGCAGCGCCAGTATAGA
AATAAGCATTCCAACTGTGATCTGGGCATGGGCGATGCCTACTACAACCAAGGCAGCGTCGGCGGCGCGGATGGTGGGCCGGTCTACGAGGAGATCCTCAGCAACCGCAACT
CGGATGTGCAGCATTACGAGGTGGGTGACTTCGATGTGGACGAGGTGTACAACAATAGCGTTGGCACTGGCGTCTTCAACAACATGAGAGCGGCGGTGGCCGCCGGCGGTAG
TCGGTACGGCGGCGGAAGCCTGAGTGGCGGCAGTGTCTCGTCCAGGAGCCAACAGCAGCAGCTCAAGAAGCAGCAGCAACAACAGTCGCTGGCTCAGCAAAGATCGGCTCGA
CGCTGCACGGCGGATGACGATGACGACGAAGACGAGGAGGAGGATGAGGAGGCAACGGCCGCGGAGCAATTGCACGACAGCGTCTGTGATGAGGATGAGGAGGAGGACGAGA
GCGACTTGGAGGATGATGCTCATGGATTGCCACCGCAGAGCGATGAGCGGATGCGTCGTCTGATGGCGATGCAGGACGAGGATTTTAAGCGGCGGTTTCAACGTCAGCTGCG
CAAACATGGAGCGCCTCTTGACTACGGGGCTTTGCCACCAGGAGCAGGACCGCAACCCGAGCACAACGGTGCGGTTTTTGGGGTTAGCGGCGGCGTTGGTGAGGGCTCCCTG
CGTGGCGCATTCCGGCAGCAGCAGCAGCAAGCACTGAATGCCAAGTCGCCAGGCGGCCGTTTGGCGGTAAATGAGCTATTCGGTCACGGCAACTCGGGACCACCGCTGCCGC
CGGCGAATCAGACGCCCGCGCAGAAGCGCCAGCAGCTACAGAAACTGTCACCACAGTCCACCACATCATCGTCGTCGCATACATCACACAGCAATCCCAATCCGCATCCCCA
CCAGCTAACCCATCCGCATCCCCATCAGCATCCTCCGCACCATCAGCAGCGCCACCTGTCGGCCATGCTCGATGAGAACAACACGGTGCGATGTTATCTGGAACCACTGGCT
AAGTGA
(SEQ ID NO: 8)

Start ATG: 1 (Reverse strand: CAT)

MLPTILSISYEHISLDLSKYQTAYACEGKKLTIECDPGDVINLIRANYGRFSITICNDHGNVEWSVNCMFPKSLSVLNSRCAHKQSCGVLAATSMFGDPCPGTHKYLEAHYQ
CISAAQTSTTTNRPSPPPWVLSNGPPIFGNGSGLIHPPGVGAGAPPPPRLPTLPGVVGISGNPGLFNVPPQHTAVTHSTPSSSTTAVGGGRLKGGATSTTTTKHPAGRHDGL
PPPPQLHHHHNHHGEDTASPTKPSSKLPAGGNATSPSNTRILTGVGGSGTDDGTLLTTKSSPNRPPGTAASGSVVPGNGSVVRTINNINLNAAGMSGGDDESKLFCGPTHAR
NLYWNMTRVGDVNVQPCPGGAAGIAKWRCVLMKRIPDSGYDEYDDDISSTTPAPSGGDCLHNSSSCEPPVSMAHKVNQRLRNFEPTWHPATPDLTQCRSLWLNNLEMRVNQR
DSSLISIANDMSEVTSSKTLYGGDMLVTTKIIQTVSEKMMHDKETFPDQRQREAMIMELLHCVVKTGSNLLDESQLSSWLDLNPEDQMRVATSLLTGLEYNAFLLADTIIRE

RSVVQKVKNILLSVRVLETKTIQSSVVFPDSDQWPLSSDRIELPRAALIDNSEGGLVRIVFAAFDRLESILKPSYDHFDLKSSRSYAILSNDSDVNAGEIQQRLRILNSKVI
SASLGKGRHIQLSQPITLTLKHLKTENVTNPTCVFWNYIDHAWSANGCSLESTNRTHSVCSCNHLTNFAILMDVVDEHQHSLFTMFDGNMRIFIYISIGICVVFIVIALLTL
KLFNGVFVKVRNGSNPLPHQRSGSRRQQNNIRDQTHESLTLTTPTSQSNVPPPSHGNTNFIQHNSIRNSHRNNLNYNVQQQQQQQQQVVAAAAAAVVLSNQPQRNMQHAMN
NLNLNLHQHGQQTAAAAAAAASIAAALQQHAVQASNASNNLNISHNYLQQQHVQQQQQQQRGQQPQPHPHRNHNLNVDGNGLDNTNNIIMQANMDDLQYKWWWCIYLCLLAI
ELLFLLGIEQTETSIFCGFITIFLHCAILSGTAWFCYEAFHSYSTLTSDELLLEVDQTPKVNCYYLLSYGLSLSVVAISLVIDPSTYTQNDYCVLMEANALFYATFVIPVLV
FFVAAIGYTFLSWIIMCRKSRTGLKTKEHTRLASVRFDIRCSFVFLLLLSAVWCSAYFYLRGAKMDDDTADVYGYCFICFNTLLGLYIFVFHCIQNEKIRREYRKYVRQHAW
LPKCLRCSKTSISSGIVTGNGPTAGTLCSVSTSKKPKLPLGVSEEAHDDPQQQQTPVPITEDAIMGATSDCELNEAQQRRTLKSGLMTGTLQAPPQTLGGHVVLERGSTLR
STGHASPTSSAGSTHLIFAHKQQQQQQQQGPLGESYYHQPDYYSWKQPSTGTGGLKTPREYYNNAGAAASSPQQAHEVFYWTQKPNSGHNGKKKRGAGGVPASPSGSLHSRT
AAASQVLFYPSYKKTKPGQPTGYPQYAEALDPPLATGNAAAYYQQQQQLRRQQLHQQQQQQQQQQLSSDEEQAEQHAHLLHLQRRAGSQQQLPAPPPHMAQYQQEFMQRQYR
NKHSNCDLGMGDAYYNQGSVGGADGGPVYEEILSNRNSDVQHYEVGDFDVDEVYNNSVGTGVFNNMRAAVAAGGSRYGGGSLSGGSVSSRSQQQQLKKQQQQQSLAQQRSAR
RCTADDDDDEDEEEDEEATAAEQLHDSVCDEDEEEDESDLEDDAHGLPPQSDERMRRLMAMQDEDFKRRFQRQLRKHGAPLDYGALPPGAGPQPEHNGAVFGVSGGVGEGSL
RGAFRQQQQQALNAKSPGGRLAVNELFGHGNSGPPLPPANQTPAQKRQQLQKLSPQSTTSSSSHTSHSNPNPHPHQLTHPHPHQHPPHHQQRHLSAMLDENNTVRCYLEPLA
K*
(SEQ ID NO: 9)

Name: Latrophilin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384743
CGCGGGCTCCATACGATTCGTATTAGCCCGCAATTGTATAAAGGTTAGTTAGTTTCTTAGCTAGTTAGTTGGACATTTAGAAGTAGGACTTTTTCTGTGCGAAAATCAGCAG
AAGTTGAAGGAAAACCCTCAGTAATTCCGCCAAAGTTCGTTGCTGGAAAAGGACTCTTTTAAGGTAATTTCATGATATTGACTCGTTTCAGTTAAGTCCCTAAAGCTAATGG
CACACTGAATACAATTAGGCATTACGGCTACGATTGCAATTGCGTTTTTTTTTTTCGATTGTTAAGTACACTTGTTTCGGCATCGGGAGAAGACAGCAAAGTTGGAAACCTA
GTTTCTGGTTCCCATTTTTATTTTTTTGTTTTTGAAGACATGAACATGCGTATCACAGTATTTGGTTGCTGCGCGTTTCTCTCTCTCTCTAACTCAAAAAAAAAGAATTTTAT
ATAAAAATGCTAAATATATATATGTATATATGTGTTATGTACGCAAAGGCGACACACATTAAGTCCTAATGTCTAAAAGGCAAGAGCGATTATTGAATTTGGTTACACCAAA
ATGGTAAAAAGGCAACGCAAGGGGCAACTAATTAGGAGCAACATTAAGTACACACAGGAAAATAGGATATGAAGCCCATCTAGGGCAATTACAAATACATTGTAGATGACCT
ATGACCTCGCGATAAAGGCCCAAACTAAAGCACTAAGAGCACATTTCCGGGGTTATATACACAGTACATAGATCGGTTTACCCCATTGTGTTCTATGGACAGGCAAAGCAAC
TAAAATCGAAATAACGATATCGTTATCTCACGATCGTTTTATGACAAACAAATTGGACCGTTTCCCCACCCATTTAAGTTGCTTTGCCTCTTTCGATACCCGAAACCGCATT
AAGTTATAGTAGAGGATTGACTTTGGCACAGCAGCTGCTGCAGGAACATTGACTGGAAGCAGGAAGCATTTCCTGCCCTGAAATGGTACCTAATCCAATTCCGTCTAGAGTA
GCGTCTCATTGCCATACTGCTGCAAGTGGGGATAGTAGCTGCCGTGCTGGCTGTACTTTTGCGAGGCATAGCTACCTGGCGGATAATACAGCCGATGCGAGGATGTGGAGTG
ATGCTGGTGATTGTGGTGGGGATGGAACTGATGATGCTGCTGCTGCTGGTGATGATGATGATGCGGCGAATGGCGCTGCTGGTGATGCACGCTCCTGCCCAGCGAATAGTTG
GACATCTGATTCAGCGATATAGAGCCGGTACCAGGCGAAGTCTTCAGTTTGTACGGACTTGGCGGCAACGGAACACTGGCCATAGAAGTTCGCACCGAATGGTAGGAGACCA
CCGTGCTATTGCAAGCCGCATGCGATCCACGCGAGCTACTATGCCCGTCCAGGTGTCCAGTTGCTGCTCCGGCAGACGTATGCTTCACCATTTTGCCATTGCTGACTTGCCT
CTTGTAGAGAGTGGAAATGCTTGGTCCAAGTACCCAGGCAAAACAGCAAATGCCCAGCATCAGTTCCATGAAGGCGCGCAACAAAAATGGCCAAAGTGGCGTATTGGTATCA
CCCGATCCCAGCCAGACATCGATGTTGGCAAACTCGTAAATGACAGCCAACAGAAGCATGGCACATGCAAGTCCATAGATGAAAGGAATATGCCAATGCCGGAACTACTGT
GTGCGCTTAGCTGTTGCAGTTGCTGTTGCACGCTGAGTGCATTGCTATTCCTCAGGATCTCCTTGGTGCGACAGTGTACCAAATAGCCGGAGATCAGATTCATCGAGCCAAA
GATCCAGTAGCAAAAAACGGGCGTAGCCACCAATATCTGTAGCGCCTTATCCGACTGATTGCCTACGAAACAGGCGCCTAGAAGAGCGTATGGTTGGGTTACTCTCCGAATT
CTCCGGCGGATCTATAGCATATCTCACCAAGCAACTCGTCAGCATCCACGAAACGGGCAACGATCACTGCTGACGTTTGAAAGGCCGGCAATCCCCAGGCCACGAAGCACAC
AAAGTTGTTTTGCGTTAGATCCTGCTGAGCCGTCTTGGCACTGTTCCTCTTCGCCGGACTGCCTCCAAAGTTCGAGGGAATGACCACATGGCCCTTGGAGTCCGGACTGTGC
CGACGAATGTCGCGGTGCCAGCCCAAACAGAGCACAGCCCACCTACAAATAGGATATGGGAATATAGTTTCTATAATATTAAGTATATATTATAATCTAATACCAAGCGCAG
GCGGCCATTCCAAAATAGTAGCGCATAAGGAAGACACTAGCGCAGGAGGCATTGGACAGTCCGTCCACGGTGAGTAGGGATTCATTGGGCGCCTGGGGATCCGTGCCGCACG
CAGTTCCCGTTCTTCCCACCATAAAGCGCACTGCCCAGCCCAAGGTGACCATGTTGTGGCACCAAATCAGCGGCGACAACAATCTGGACCACTTGGCGCTGGCCAAACGACT
GCCATCGCTGGCCAAAAGACACACGGTGGCTACCAGGGCCAAACCCAAAGCGGCATAGGCCCATGTGGAGACCCAGATCTCTGCCAGATGCTTCTCCGCCGGTGTGAACAGT
ATGTCCGCCTCGCAGAGTGGTGCACAGCGTCCAGATCTGGGTAGCCTTACGTACAGATGCGATTTGGCCAGTCCCGAGCAATCCATGGGCAGTTTACCGCCCAATCCACCGG
GTATTCCTTGGCCCGGTAGCCCATACAAATCCTGCTCCTGCTGCGGCTGATGCAATTCACCGGGACCTTCCATACACATCGTCTCGTGGTTGTTCTCTCGCGGAAATTTATC
GCAGTCGAGAGCCGGTGGCCATGGAAATCCGAATCCCTGCAAAACTGGATGACAACGTATTCGCACGGACTCACAGAGACTCCGGCACGGACCGATGGCATGGACTGGAGCT
TTGGGCGTGCACATGGGCACGTAGGCGGCGCACAGGAACAGCTTCAACTGGGAACTGCAGTCGTATTCGATCAGTGGTGCAAAGGTTTGCAGTGTGTACTCCACGTCCGTTT
GCATTTCGTTACCCACCAGATTTGGCATGGAGGTCTCATTGTAGCCGATCTTGCGGCACATCTCAATGCGAATGGTTTCGCATTGCCGGAAGGCGGGGATTTCTGGCGGAGA
ACTGGAACTGGAACTGGCACTGGGGTTGCCGCTCGTTGAGGATTTTGAGATGCGCGGATGCAGGAGGATCACCACCAGGAGACACAATATGCAGGTTGGCTTCATTGCGGCT
GATTTGTTGTTCTTTCGTTCCGAAGAATACGAATTCGATACGTTTCTTCGCTTGGTGTTCCGATTTCCGTTTCTCTAAAGGAATCGAAATGTGGAAAAGATCAGATTTCGGG
GTTTTCTTCGCGCTGGGTTCCCTCTTTTCTTTCCAAAGAATCCGAACTATATGCAGATTCTTCGTTCTTCGGTGGAACGAACGCTGGAAGGGGTAACCTGTTCCAGGTGCGT
TCGCCCTCTCTCGCTCGCACCCTCTCACTCGTTCCAACAACTTCCAGCTCTCCTTCTTTCCAGCTTGTTATTGTTGCTTTTGGCTTGTCAACTCGTTAACTACCTCTTTTTT
TTCGGCGTCGTTCTTGTTATCGCTTCTCCTCCTTCCCCCACTCGTTCTTCAGCCCCAACCACTCCCCTTTCTCCACCGCACGCCGCCCAACCCGCTGATCGACGTCGATAAA
CGGGTTTTTCCTCAGGGATTAACCCGAAAACGATCTAGCAAACCATCGAATCTTTCACACATCTCAGTGAACGAATAGTGTATTTGTTCGATGTGCTTAAACGAATGGCGC
AGTCAGTTTGTAGCCTTAGCGAATGTATTCTAATCTAATTGGTATTAACATCGAACGTTTATATATTTTATTTTGTTTTCGTGGTAAAATTCAACGCATACTGGTCACTCTG
AGCTAGAGATGGCTCTGCGTCACGTCACCTCGCATCGATTGCAATTGCTCAACGGATCTGCTATCGATATCAACAAGTGTGTTGGTTACTAAATGCTTTTAGATCAAAAGGC
GCCCCCTGGCGGAGTTATGCGTAATCAAAGAACATTGAGAAGGGCGCGCTTTTCAAAAACAAATTTGTCCAAATAATAAGAAACATTTAAGCATAATAAGAACATAAATGTT
CAATAAATTCTATAGGTGCAGTAAAATACGTGTTTTGCTAGCACAGCTATTTTAAGCTGAGAACTTACAATATTTTAGTACATTGTGCTTAAGGTTTAAAGATACATATATT
GAAAAAGTTCTGGAAATGCTTTGAAGAACAAAGCTTTAAGAGGCGCGCTTTCCGAGTATGTGAGCTTTTTTCAAATGGCACACTTTCTGCGGAATATAAAATAAAAAAACC
TAATAATATGCCGCAAGACCACTTAAGATAAAATATTAAAAATATTAAAATTGATTATATTTTGTGCACCAAAAGTGGCGGTAGCTTAATATCGCAAAACCAAATTAATAGC
AATTCTGAGCCAAACCGCGATTGCCATTTCTGCATGTTGGTCCTCCGTCCGTTTCGCCATCTCCCTTTGGCCACCGCTTGGCCTCTCTTTCGCACACACATGGCAAACACAA
GCCACGCCCATGTATTTCAACCCCCTTGACAAAAATACGTAACATGCCACTTTTATGTATGAAACGTGACAAATACGCACATGTGTGAGAAAGAGATCCACCGAGCTGGCGA
GAAAGAGAGAGAGAGACGCTGGACGTTTATGTGTTTGGCAGGACTCGCGTATGCAGTGTGTGTGTGCGTGTGTGTGTGAGTGTGTGTGAGTGTTTGTTGTTGGTTAACGCAC
ACGCGCAGACGTAGCAACACATACACATATGTGGAAAGGATATGTACATACATACATACATACATACATATGCCCGCTGACTTTTTCGCCACTGCCAA
(SEQ ID NO: 10)

Exon: 3914..3576
Exon: 3322..2232
Exon: 2170..1932
Exon: 1854..1001
Start ATG: 3241 (Reverse strand: CAT)

Transcript No. : CT9057
GACCAGTATCCGTTGAATTTTACCACGAAAACAAAATAAAATATATAAACGTTCGATGTTAATACCAATTAGATTAGAATACATTCGCTAAGGCTACAAACTGACTGCGCCA
TTCGTTTAAGCACATCGAAACAAATACACTATTCGTTCACTGAGATGTGTGAAAGATTCGATGGTTTGCTAGATCGTTTTCGGGTTAATCCCTGAGGAAAAACCCGTTTATC
GACGTCGATCAGCGGGTTGGGCGGCGTGCGGTGGAGAAAGGGGAGTGGTTGGGGCTGAAGAACGAGTGGGGGAAGGAGGAGAAGCGATAACAAGAACGACGCCGAAAAAAAA

GAGAGAAACGGAAATCGGAACACCAAGCGAAGAAACGTATCGAATTCGTATTCTTCGGAACGAAAGAACAACAAATCAGCCGCAATGAAGCCAACCTGCATATTGTGTCTCC
TGGTGGTGATCCTCCTGCATCCGCGCATCTCAAAATCCTCAACGAGCGGCAACCCCAGTGCCAGTTCCAGTTCCAGTTCTCCGCCAGAAATCCCCGCCTTCCGGCAATGCGA
AACCATTCGCATTGAGATGTGCCGCAAGATCGGCTACAATGAGACCTCCATGCCAAATCTGGTGGGTAACGAAATGCAAACGGACGTGGAGTACACACTGCAAACCTTTGCA
CCACTGATCGAATACGACTGCAGTTCCCAGTTGAAGCTGTTCCTGTGCGCCGCCTACGTGCCCATGTGCACGCCCAAAGCTCCAGTCCATGCCATCGGTCCGTGCCGGAGTC
TCTGTGAGTCCGTGCGAATACGTTGTCATCCAGTTTTGCAGGGATTCGGATTTCCATGGCCACCGGCTCTCGACTGCGATAAATTTCCGCGAGAGAACAACCACGAGACGAT
GTGTATGGAAGGTCCCGGTGAATTGCATCAGCCGCAGCAGGAGCAGGATTTGTATGGGCTACCGGGCCAAGGAATACCCGGTGGATTGGGCGGTAAACTGCCCATGGATTGC
TCGGGACTGGCCAAATCGCATCTGTACGTAAGGCTACCCAGATCTGGACGCTGTGCACCACTCTGCGAGGCGGACATACTGTTCACACCGGCGGAGAAGCATCTGGCAGAGA
TCTGGGTCTCCACATGGGCCTATGCCGCTTTGGGTTTGGCCCTGGTAGCCACCGTGTGTCTTTTGGCCAGCGATGGCAGTCGTTTGGCCAGCGCCAAGTGGTCCAGATTGTT
GTCGCCGCTGATTTGGTGCCACAACATGGTCACCTTGGGCTGGGCAGTGCGCTTTATGGTGGGAAGAACGGAACTGCGTGCGGCACGGATCCCCAGGCGCCCAATGAATCC
CTACTCACCGTGGACGGACTGTCCAATGCCTCCTGCGCTAGTGTCTTCCTTATGCGCTACTATTTGGAATGGCCGCCTGCGCTTGGTGGGCTGTGCTCTGTTTGGGCTGGC
ACCGCGACATTCGTCGGCACAGTCCGGACTCCAAGGGCCATGTGGTCATTCCCTCGAACTTTGGAGGCAGTCCGGCGAAGAGGAACAGTGCCAAGACGGCTCAGCAGGATCT
AACGCAAAACAACTTTGTGTGCTTCGTGGCCTGGGGATTGCCGGCCTTTCAAACGTCAGCAGTGATCGTTGCCCGTTTCGTGGATGCTGACGAGTTGCTTGGCAATCAGTCG
GATAAGGCGCTACAGATATTGGTGGCTACGCCCGTTTTTGCTACTGGATCTTTGGCTCGATGAATCTGATCTCCGGCTATTTGGTACACTGTCGCACCAAGGAGATCCTGA
GGAATAGCAATGCACTCAGCGTGCAACAGCAACTGCAACAGCTAAGCGCACACAGTAGTTCCGGCATTGGCATATTCCTTTTCATCTATGGACTTGCATGTGCCATGCTTCT
GTTGGCTGTCATTTACGAGTTTGCCAACATCGATGTCTGGCTGGGATCGGGTGATACCAATACGCCACTTTGGCCATTTTTGTTGCGCGCCTTCATGGAACTGATGCTGGGC
ATTTGCTGTTTTGCCTGGGTACTTGGACCAAGCATTTCCACTCTCTACAAGAGGCAAGTCAGCAATGGCAAAATGGTGAAGCATACGTCTGCCGGAGCAGCAACTGGACACC
TGGACGGGCATAGTAGCTCGCGTGGATCGCATGCGGCTTGCAATAGCACGGTGGTCTCCTACCATTCGGTGCGAACTTCTATGGCCAGTGTTCCGTTGCCGCCAAGTCCGTA
CAAACTGAAGACTTCGCCTGGTACCGGCTCTATATCGCTGAATCAGATGTCCAACTATTCGCTGGGCAGGAGCGTGCATCACCAGCAGCGCCATTCGCCGCATCATCATCAT
CACCAGCAGCAGCAGCATCATCAGTTCCATCCCCACCACAATCACCAGCATCACTCCACATCCTCGCATCGGCTGTATTATCCGCCAGGTAGCTATGCCTCGCAAAAGTACA
GCCAGCACGGCAGCTACTATCCCCACTTGCAGCAGTATGGCAATGAGACGCTACTCTAG
(SEQ ID NO: 11)

Start ATG: 421 (Reverse strand: CAT)

MKPTCILCLLVVILLHPRISKSSTSGNPSASSSSSSPPEIPAFRQCETIRIEMCRKIGYNETSMPNLVGNEMQTDVEYTLQTFAPLIEYDCSSQLKLFLCAAYVPMCTPKAP
VHAIGPCRSLCESVRIRCHPVLQGFGFPWPPALDCDKFPRENNHETMCMEGPGELHQPQQEQDLYGLPGQGIPGGLGGKLPMDCSGLAKSHLYVRLPRSGRCAPLCEADILF
TPAEKHLAEIWVSTWAYAALGLALVATVCLLASDGSRLASAKWSRLLSPLIWCHNMVTLGWAVRFMVGRTGTACGTDPQAPNESLLTVDGLSNASCASVFLMRYYFGMAACA
WWAVLCLGWHRDIRRHSPDSKGHVVIPSNFGGSPAKRNSAKTAQQDLTQNNFVCFVAWGLPAFQTSAVIVARFVDADELLGNQSDKALQILVATPVFCYWIFGSMNLISGYL
VHCRTKEILRNSNALSVQQQLQQLSAHSSSGIGIFLFIYGLACAMLLLAVIYEFANIDVWLGSGDTNTPLWPFLLRAFMELMLGICCFAWVLGPSISTLYKRQVSNGKMVKH
TSAGAATGHLDGHSSSRGSHAACNSTVVSYHSVRTSMASVPLPPSPYKLKTSPGTGSISLNQMSNYSLGRSVHHQQRHSPHHHHHQQQQHHQFHPHHNHQHHSTSSHRLYYP
PGSYASQKYSQHGSYYPHLQQYGNETLL*
(SEQ ID NO: 12)

Name: Frizzled 4 -like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384825
TCATACTCACTGGCTGCCAACCCCTACTCTACTCTGACTACTTTGGTTCTGGTTTTGGATCTTGGCTGATGCTAGGTTCTAACTTTAAGTACGATTCTTAGTATTGCTCACT
GGATACACTGTAGATACACCGCAGATACAGCCGTCCTGCGGCGGCACAACAACAAGAAAACAAACGCTAATGCAGCTACTTACGACTAGGAGGATGATTAGTGGCAAGGGGT
TAAGGAGCGGGTTAATCACAACTGGCTGGGAAACTGAAAGTGGCCGACAAGCGCATCTACTTAACCTAGATGATGGTCCTACGCGGCCCAGGGTCAATCACATTGCGGAGTA
GGTTCTACACAGAGAGATGCGATTTGGTCAAAAACTAAGCTTAAACGACCGCGATGTTATCTTATCAATTAAAACGAAAAAAATGGAATCAACATTTTTATAGACCCATTTA
AAGAGCTATCATCTTTTTTCACCGAAATACTTAAGGAAATAATATGCATGCTTAGTAAAGGAACACTACACTAGGTCTTAAGAAATGATTATCAAGAAGTAAACATTTGGCA
CCCCGAGGCGAATACTGATACTGGTAGTTTTTAGAAAACTTTGCTGATACCGAAAACTCGCGAATTAAGATAAATTCTAACATTGCACAAATTTCTTAGTTTACATTGTAAA
TTTTGTCTACTTCTAGGCAGAGCACAATGTGCCTGCAATTTTCTCCGTGTTGTTGATTCTGCTAAGTTCAGGAAAATGACTAGCTATGTGAACTTGACTTGAGGAGGGAGGG
GGAGGTTCGTATCGTGACTTATCTAGCGTTTTACCCTAATGCATTTGATTGCTTTAAGCTGATGATTGGTCAGTTGCACACGCGCTATTCTGGGGCAGATTTTGTGGGCGTG
TCGTAAAATAATTGCTTGGTGTTTCGCTTCTAATTGACGAGGACTTGGACTTGGATTTGAAGTGTGGACCGTGTACAGAGTGGACAATGGACAGGGCATCTTATTTATTCGG
TGCCATTCGCAGGCGCCTGCTGCTCCTCTTCCTGGGAATCGCCGAGGGAAAGCCTCGACTTGCCCGCTGTTTCCATCGATTTCCAGGTGCCTCCACTGGTGGCCTGCTGCAG
CGAAATGCGAGGTATGGACTTGCAGCTGCCCAAGATAGAGGTGCTATCTCCGGGACGGTCCTCGGATTCGGCCAGGGCGCTGAGACTGCAAGTGGGCGGCAGGATGCGGTGT
GGAAAGCGTGGCGAACCCCTTACCCTTTCCCTTAAACTGGGTGGCTGCTGCACCACCGAAGGTCTCCGGTTGCTGGGCAATTGGAAGTTCACCTTGGTCTCCAGCAGTTTCA
CATACTTTCCACCCAACGTTTGGTCCCCCTGTTCACCGCCAGAAATAGTGGGGATAGGCAGGAAATTCGGATCCTTGGTAAGGTGATCCAGGTGAGTACCGCTTTTGGTCAA
TGAACCCCTATGGATGGGCTTTAGGTGGTTACCATTCGGGTGGGAGGACGCCTCTGAGGGAATGGAATCCTGATTTGGATTATTGTATATGGCATTGCAGGACGCCGTTTGA
GAGGAGGGAATCAGGCGGGTGAACAGTCCCATGTGCGATTTAAAGGCCGCCTTCAGACCCTTTTGGTCCGCCAAGCTGGCGGTGGATCCCTTGTTGCTCGAAGTGGTGCGAG
AAAATATCCCTCGTAATCGGTTTATCAAGCCCGTCTTAGCCTCCGTGGCCTGCTGGTGGGCTAAACGGAGGTTAATCAAGTCCTGTGCAGGCAGGCAGGCCTCCTCTAATTG
GGGCTGGGAGGCAAAGGTAACAGAACGCCTCATCGGGATCTGCATGTACTCTGCTGACGGCCAACTGGCCCGTGGAATCACAGGAGGCAGCACGGAAAGCAGGAGATTGGGC
AGTATTCCCGAGATACTTGGAGTTCTGTTTTTACTTGCAGCCGGAGTCTGCGAGATGCCACTGGTTACCGTCAGTTCTGGCTTTAGATGACCTCCCCCAGTTAAGAGTGATG
TGGACGACGAAGTTGTGGTGGAGGTGGTGGTTAGTCCCGACTCGAGTAGCCTCTCCAGCTTTCGAATCTCGGCGTCCAGAGCCTGGATTTCCTTCTTATAGACCCTGTTTTG
CACCTCCACTCGATACTGCAGTTCTCGGCGATCATCGACCACGAATCGGCGTGTGTTGCACTCCATCTTAAGGCCCATACTGTGGATAACCGGATCGATAATATCGTCTGAG
GGAGAAGGAGAGTTCCCAAAGGATTAGGACACAACTTTATAAAGGTTCTAGAAATTAGGAGACTAGAACTTAAAAGTATTTCCTAAGTGAGCAAGATCTTATTGCACCTTCA
AGCGCTACTTACTTCTTGCCCAAATATCATGGAGTTTTGGGATGAAAAGCAGACAAAGGGTTGCAGTGGTGCTGGTTAAAATCAGAGCTGTGATTGTGATGAAGGCCAGGGT
GACTCGCTCCGAAATCAAGTTGGCCAGCACCACGACGATGGCGCTGGTGATGACCACACTGTATACAGACACTCCGATGTACTGCGAGTCATTGAGAGCAGGTATTTTTACG
TGGCGCGTCTCCCAGGCCATATAGACACCCACCACAAGAAGAAGGCCTTTGTAGGCGTACAGGACACTCAACCACGTTTGCGTGTGCTGCGAACGGCAAACTTCAACCTAAA
ATAGATAGATAGATAGATATTCATAGATTCTAAATGAATTAATACTTTTTCAACATAATTAAAAGTAATCTGATAATTTGAAGCATATTAATATTTTTCTTACCTGAGGCTG
GTAAACGACACTTCTATCAGTCGCACTGATCTCGAGCGTCAGGTTGTGAAGATGGCGCTCCATTGGATCGGTGACCACCCAAAGGGTTACGAGCAGCGCATCCACCAGAAGC
AATCCGCCGACGAGCAAGATCAGTTGAATGTCCTGCAGCATCTTGTCCTTGAAAACGCTGCCGGTACGAGTGAATATCCGATGCACTCTGTAGGTCTTGGCAAACATCGATC
CAAAGGCCAACGAGAATCCGGCGGAGAGCAGATAGACGCGGGCCTGGTGAGCATGATAAGCGGATGACGGGGAAGTTCAGGTTAGGGAGAAAGTACTGTAGTAGCAGCAGCA
AAGTAACACACTAAGTGTCATGTCAGGCAGACCCAAAGAAAACAAGCCGGCGGCAACAATGGGGGAAACAAGTCATCAATGGACCACAATTCTCAGAGTGCTTGAGACTTTG
TTGTGCGTTTCAAATTTATGCAAATAAAGGAATATCACAAAAGTTGTAAATTTCCGTGTCATGTGTTTGCATTTCCCTGCCACAAGCCTCAACAAAACAAAAACAAGCAGGC
AACTGGAAAAGCGGAACGACAGTAGAAGTGGTTACAGAAACAGAAACAACAGCGACCAGCAAATTTCCCAGGAGTGGCGTTCGAGGCAGTTGTTAAGGGGGTCAAAGTGGAA
AGGCGCAGATCGGGATAGAAGAAAGCTCGTCAAACTTGACTTGACTTCTGGCTTAAAGAGCAGCAGATTTGCAGCGAAACCGAACGGCTATTTATAGCCAGCAGGGGCAGTC
GTGTATCTTGAAAAGACACGCGGAATAATCAATATTCAGTTTATATAATCCACTTTAAAATAATAGAAATTAACAAAAAATTGCCACAATTATCAGTAATTTTTATACTTCG
CAGGCTATTTTATCGCCATCTTCGTGTATACGAATACGACAATTAGTTCAAGCAATATTTCTTTCAGTGAATAAATAAAGCAAACACATGAGCGCGACTCGAAAGTTACCAG
CCTGAAACTGAAAGTCATATTTGATGAGGACCCCACATGCTTGCTTACACACAGAAGCCCGTCAAAGTCAAACAAAGACGGCTGAGAGGAAAGGCGGATGAGAATGCGGATT
CGGATAGGAATGGCGATAGGGATGGGGATGTCGGAGCCGAAGGATAACAACTCACCGTGCAGACCGTTGCGAAAGAGTCCTCCGCCGAGGGCAGCGTCGAGTGGTCCAAGCC
CAAAAGGATGACGGTGGCGTACACAAAGATGCAGCCCACTGCGGTGATGTTGCTCAGCTTCGGGCTGGAAAGTTTAATTGCCCTGTGGAGTGAGAAGGGGGATGGTCATGGT
TATAAGCTGAGGCGCCGCAATCCTTTGCAGAGAAAGTGCAGAGTAACCGTAGAATGCGAAACATATATACGAACTCAACTAACGATGGGTCAGAAATACAGCTGGAAAGCCA

AGCAATTACTTCACTTCAAAAACCGAAACAAGTCACATGTGTAATGGGTTCAATGGGTTGAGCACCGAAAGCAGTTAACTCATCATGGAATATAAAGAAACAGCCCGCACAA
ATTCTCTTTTATATGGCCACTAAATATTTTATAGTCGTTTAATGAAAAATAATATATTCAAACAATCAGGGCCCATAAAAAGAGTAAGCCAATAATTTATGTACTTATCGCC
TTACTTCAGCTTCCGAAAGTGCAGATTGAACGCCAGGAAGGCGATGGCCAGAGCGATTCCCACGCTGGAGAGGGTGGCGATGGTGTAGAAGGCCAGTGGAGCGATGGTCGCC
ACCCGCAGCTTGAACACCCGCTTGGCGATGGGTACCTGCCCGCTGTGCCACTTCACCGGCCGGCAGCGGGGACACCGGAAGTCCAGGGCATCCGTGGCCGGATAGTAGAGGG
CCACCGGTTCCAGCAAACCGCGCTGGATTTGATAGAAGGCAGTGGTGCCAACGCGATCTGGGCCGCTGAAGGAAACGGGGCCCTACCAAAAAGAGAAATACTCGTGATGAAC
TAGATTTAAGAACCTTAGGATCAAAGTTAAATGGGTTGGATGATTTCAGTTGAAAGAGATAACTCACCGACACTCCCAGGAAGTGGAGCTTGCCCATTTGCTGCAGGAACTC
CCAGGCCATGTCGCTGCGGGTGTAATCGAATCCGTCCAGCTTCGACTGCTCCTCGTTCCGCCGCCAGTGCTCCTCAGCGGCTCTCAAGGCCAGGGCGATGGCCCACACGGCG
TCGTAGGTTTGCGGCGCGTACTGCGAGATCGCCTCCGGAAAGAGGTGCCCTCCGCTGGTGCCTCCCACGCCCCTTCTCCTCCGCCGACGAGAGTCAGATTGCGTTGCAGCGA
TACTGATCCTGGAACCATAACCGGAGCCAAATCCATCCTGGCCGTGGAACTGGGCGGATTGCTTGCGCAGCTGGGAGTTGAACATGTGATTGTTCTGCGAAGTGGAAACGCG
ATTAAACAAGCCTCGAATGCCGCAGAAATGTCGCCAGTAATGAGATTGCCTATGGCAGCTACCACTCCTGCCCCAATTTCATGACCAGCACTAAAAGAGCCATAAAAACGCG
TCTGCGATGGTGAATGCGAATGCGATGACAGGCCGGACTTAGGGGGCTTCGCACTTGGCAATCACCATCGCCATCGCCATTGCTGAAACCATCGCGAAACGCAATCGCAGCC
TCGGGATTGTCTAATCAAATAAATCATTTCATTTCGGACACGCACAGCAAGACGGACAAGACAGGACAACGAGAAGGACAAGATAGTCGTCTCTGAATAAAAAAAAAAAAAA
AAAAAAACGGGGAATAAAAGTTTCGGCAATTCTGGAAAATGTTGCTTGAAGCTGCCTTTGAGAGTTGGGTTCAATCCAATTTGGATCCCGAAATTTATATGCATGTATAACT
TAACAACAATCTATTCTTGAAATTCATCTTTACGTTAATCTCATACTGTCATGAAATTAATTTATTTATTTTCCCTTTAATATTGAATGATATTTTCTCGCAGTGCACCTCC
TCAACTTACCAGTCCACTATAGCTGACGTTATTTCCAACGATGCTGTTGTGCGTTGAGACCACGATGAGGTTCTCGACGGCCAGCTGCAGTTCGTGGTTAGAGCAGGCGGTG
CGCTGGTCCGGCCACCACGGAGCCCCCATGCTCTCGTGGAGGATCCAGGCGTAGTCCGCCCCGAACATTCGAAGCCTGTAGGCCTCGCACAGGATCTGGGGGGCCAGCTCCT
GCGAGAAGCTGCCGATGATGATGCGCGTGTCCGTCTCCTGTGGCGATACGAACGAAACGGAACGGCATGTGCCCATTTGTCATTTGTCATCCTTCGATTAGCACAAGTGCCA
ATCCACACTCAAAAAAAATGGGTAACAATCATCATGCTGATAGAGGGTCTCTAGGCGATGTGGTCTGCTAATGAAACATTTATCCATGCGGACTGGACCAACAAAAAACTTG
AAGATAAACTCAGAGAAATAAGTTAACGTAAAAAACGTAATTATAATCATACATATGAACCTTTTTTTAAAACTTACTAAATTAAGTTGCCAGGCAAATAAATTGGTCTA
TTTAAATATTCCCCAAATACGTTCATTTTCATATTTGTTCATGCATAAGCTTTTAATATAAACAGAACAGCGCTAGCACCATTTTAATTTAAACATTTTCGAAAATTGAAAT
TTAGTTAAATTTAATGTTGGTTTTCAAACAAATATTCATAAATGTTTACGTTTATACGTTTGCGTATACGAGTATAAAATAAAACGAAAATGTGTATTTAAATGGCCATTT
GTTGCTCACCCTAAGTAGCAGCAGCTGCTCCTTGAAGTCGGTGGCCGCAAAGGTGATGGTGGCGGCACAGGATATGTTGGCCGCCTCCAGTTCGGTGACCAGGTTGTTCACC
GCCAGCGAGTGAACCTCCTCGTTCTGCGAGAAAGTGGTCACCGTGCCCCAGCCAAACTTCCGAATGAAAGCGATGCGCGCCGGATTGTGTGAGGAGTCCGGGGCCACCGTCC
TGTAGAAGTAGGGGAACTCCCGCCTGTCGCTCAACGCCGGCGATGTGGAACCGAAGGATACCTTTCATCGAACGAAATCGTTAGACTATCTCTACCGTACATGGGGTCGCTT
AAAATGCAAGACACGCATCAAGAAATCCAGCCCTCTAGGTCATGTGTCTTCCCAGTTTCCCCACTCATTTAACACACGCACCGCCGAAATGAAACTTCAAAGGCCAAATGCA
AATGCTCTAAGTGAATTTGTGGACCTCATAATTTGTGGGCTGTGGAAATGTTTATGGCTCAAAGGTTTATGGGGTGCTGGAATCGGCTGGATTCCGCAGCATTTATGCTAAT
CGAGCTACGCACGTGAGTTGATTAGATTTTTGAGAAATTGCCCGATACGAGGAGTTGGTTGGATGAACCGAGCGAGCTGGATTATTTGCATTTTAATCAGTCGCAGATGACT
CACCTGCACGATGTTCCAGTAGGGCACCACCTTCGCCAGGCTCTCGGTGACCTCCGAGCAGGCCGATCCCAGCAGCATCACCATCCTCGTCGAGGGCTGTGTGTAGATGGCG
TGGAAGAAGCGATCCACGCCCACTCCAGGATCACACTGATTGTCGGAGAAGAAATGTGTGAAGAAACGGAGGAGCAAAGGACACCACATCAGAGATGAGTAATAGAGGGGGC
AGGGAAAATAGCAGTGGCAGAGGCCGTTGCTCGCTGACTCGTTTCGACGCAATAAAAAGGCAACATGATGGCGCAATAAACATTTTTAAAATGTCGACTTTTAGAGAGGCAC
AGCCGGAGATGCGGCTCCAGATATGACAGTCTGTTTAAGGTCGATGTGCATATGCATACACTTAAAAAAATCGTTTGAAAACCTAGTTCACTAAGTGCTGACATCACAGTCG
AGGTATTACGTATTGACGGTTTAATTTATCAGAAACTAAACTACTATTATGCATTTTCTTTTTTTTACCATCAATTAATTGTAGCTTTAATACACTACTGCACCAAGTTAAA
AATGATTGCAGGTAAATAGTTTACATTGTTGGTAACTTAGTCAGTGGCATTGCTTTTCTTCAGTGACAGTGTTCTGTACATTTGCGTGCCCCATCAAAACCGGCAACGACTC
ATCGCAATAAAAGCGCGAAAGACGGAAGGTTGATGACATATTTGCCTCTCTTTTAGGGCCACAGCCGCAGCTGCTATCCGCACTGTTCTTGTTGTTGCAGATTTTCTTATTA
AGTTTTAATGTTGCACGCGCTACTGAATGAAAATTCAATTTCGTAATGGAACAGCGACGCCTAGAAGTGGCGCCCAACGTGGGGCATACGGTTACGCCACGTAGAAATGACA
TTCTCACCTGAGTATCGTTGGTCACGAGCTCGAGGGTGTAGCCCGGCAGCAGGCGCTTGCGGTTGATGTGCTCCACGGCCATGGTGGCAGCTCCCAATTCGCTCAGACCATC
CGGACGTGGTCCCCGCGATGTGGACAGCTCAAAGAGTCCGAGCAGCACGATTTTGCCATGATTGCGCTGCATTTCGGCCGGCGACATGCGGCGTTCGCTTTCGAACACCTGG
TTGACCTTCAAGTATCGCTGCCTGGGGGGCAGATTGAGGAGGTGGTGGGCGTGGGAGGCATTGGTGCTTCCGGGCAGATTGCTGTGCGAATTGAGGCGCTGCTGCCGGCGTT
TGACTTTCGTCCGTTTGTACTTCCAATCGGAGCGCGGTGGAATCGGAGTCCCCCCTGCAGTCACAGTCGATGCCGATTCTCCGGGTGGCGAGCTGGATGGGCTGGATGTGCT
AATTGAGGCAGCACCTGGCCTAATTGCCTCGTAGCCCAGGGCCTGCAGCTCGGCTGATGACTCCATGGCAGCCGCTGCCGACGTTGCCCAGGCAAGGGCCAGGACTAGCCTC
AGTCCCACGGCCGGCCATGGACCGTATCTGGTCCCTTGGACCGGTTGAATTATGCGCATCCCGTCGGCAATGGCCAATGGCCCATTAATCACCAAGGCATTTGAATCGTCGG
CTATCAGAGGCTGCAGCAACAAACAAACACGAGCCACACAGCACAGGACACCGGAATATATATCGCCTTTTTATGCACTCTGCCGACAGGATACGCTTATCGTTTCTCTTCC
GACGAGCGGCTCTTCCTTCTGCTCCCTCCTCTTGTATCGGCGTCAATTTTATCGCCTTTTTATTGGCCTCACTTTTGCTCGCATTTTGTATCTTTTTGCACATTCACTCTGC
ACAGCAGTCGACCATAAAAGTATGCAATGCAAAGGACTTTCTCGATAGGTCCGATTGACAGGATTGGGTAACACGGAAAAAACTACACGCCATTCGCTAACTTTCAGTCAC
TTTTCGTAGAAAAAAATGCTTATTTTTATTATGCACTGTTTATTAAGGATTAACACAATTTAGGTGAATATTTATTAGTATTTTGCTGTAGTATGTGATTTTATTAAAATGT
AATTTGAATCTACATATGTATATATATGTATATATGTGAAAACATTGATTTTTTCATGAGAATGTGAAAGGGAATGTTAATGGAAATACCGAAACAAATGTCTAGGGAGGCA
TAAAAAGGCATTTCACACCAATTTGGCAGTCGAAAGGATTATGACGAAGGCTTCGATGATATGGACACACCGCGCCGCTCACGTCTCGTCCCATGACTTGTTATGTGCCTGC
CACTTGGCCATTATTTTTTATGCTAATGGCCAAACATTTTGCGCTTTCATTGCGGGCTTTTGGACTACTTTTTGTCGAGTCTCACGCGCCACAGATTTCTGTGGCGAACGTT
TTTCTCTGTATGCGCGCGTATGCATAATTGTTTCAAGATTACCAACTTGCAGTCCCTTTGCCGCTTTCCCGCTTTTCCACCGCCCACTTGCCTGCACCTAAGCCCACTTCAT
AGACGCTTTCCTTCGCGCTTTCTGCTTTTCCCACAGCACGCTTTTCGCTGT
(SEQ ID NO: 13)

Exon: 8795..8184
Exon: 7427..7284
Exon: 6893..6618
Exon: 6085..5834
Exon: 5246..4884
Exon: 4786..4485
Exon: 4114..3976
Exon: 3067..2792
Exon: 2683..2365
Exon: 2235..1001
Start ATG: 8795 (Reverse strand: CAT)

Transcript No. : CT9836
ATGCGCATAATTCAACCGGTCCAAGGGACCAGATACGGTCCATGGCCGGCCGTGGGACTGAGGCTAGTCCTGGCCCTTGCCTGGGCAACGTCGGCAGCGGCTGCCATGGAGT
CATCAGCCGAGCTGCAGGCCCTGGGCTACGAGGCAATTAGGCCAGGTGCTGCCTCAATTAGCACATCCAGCCCATCCAGCTCGCCACCCGGAGAATCGGCATCGACTGTGAC
TGCAGGGGGGACTCCGATTCCACCGCGCTCCGATTGGAAGTACAAACGGACGAAAGTCAAACGCCGGCAGCAGCGCCTCAATTCGCACAGCAATCTGCCCGGAAGCACCAAT
GCCTCCCACGCCCACCACCTCCTCAATCTGCCCCCCAGGCAGCGATACTTGAAGGTCAACCAGGTGTTCGAAAGCGAACGCCGCATGTCGCCGGCCGAAATGCAGCGCAATC
ATGGCAAAATCGTGCTGCTCGGACTCTTTGAGCTGTCCACATCGCGGGGACCACGTCCGGATGGTCTGAGCGAATTGGGAGCTGCCACCATGGCCGTGGAGCACATCAACCG
CAAGCGCCTGCTGCCGGGCTACACCCTCGAGCTCGTGACCAACGATACTCAGTGTGATCCTGGAGTGGGCGTGGATCGCTTCTTCCACGCCATCTACACACAGCCCTCGACG
AGGATGGTGATGCTGCTGGGATCGGCCTGCTCGGAGGTCACCGAGAGCCTGGCGAAGGTGGTGCCCTACTGGAACATCGTGCAGGTATCCTTCGGTTCCACATCGCCGGCGT
TGAGCGACAGGCGGGAGTTCCCCTACTTCTACAGGACGGTGGCCCCGGACTCCTCACACAATCCGGCGCGCATCGCTTTCATTCGGAAGTTTGGCTGGGGCACGGTGACCAC
TTTCTCGCAGAACGAGGAGGTTCACTCGCTGGCGGTGAACAACCTGGTCACCGAACTGGAGGCGGCCAACATATCCTGTGCCGCCACCATCACCTTTGCGGCCACCGACTTC
AAGGAGCAGCTGCTGCTACTTAGGGAGACGGACACGCGCATCATCATCGGCAGCTTCTCGCAGGAGCTGGCCCCCCAGATCCTGTGCGAGGCCTACAGGCTTCGAATGTTCG
GGGCGGACTACGCCTGGATCCTCCACGAGAGCATGGGGGCTCCGTGGTGGCCGGACCAGCGCACCGCCTGCTCTAACCACGAACTGCAGCTGGCCGTCGAGAACCTCATCGT

GGTCTCAACGCACAACAGCATCGTTGGAAATAACGTCAGCTATAGTGGACTGAACAATCACATGTTCAACTCCCAGCTGCGCAAGCAATCCGCCCAGTTCCACGGCCAGGAT
GGATTTGGCTCCGGTTATGGTTCCAGGATCAGTATCGCTGCAACGCAATCTGACTCTCGTCGGCGGAGGAGAAGGGGCGTGGGAGGCACCAGCGGAGGGCACCTCTTTCCGG
AGGCGATCTCGCAGTACGCGCCGCAAACCTACGACGCCGTGTGGGCCATCGCCCTGGCCTTGAGAGCCGCTGAGGAGCACTGGCGGCGGAACGAGGAGCAGTCGAAGCTGGA
CGGATTCGATTACACCCGCAGCGACATGGCCTGGGAGTTCCTGCAGCAAATGGGCAAGCTCCACTTCCTGGGAGTGTCGGGCCCCGTTTCCTTCAGCGGCCCAGATCGCGTT
GGCACCACTGCCTTCTATCAAATCCAGCGCGGTTTGCTGGAACCGGTGGCCCTCTACTATCCGGCCACGGATGCCCTGGACTTCCGGTGTCCCCGCTGCCGGCCGGTGAAGT
GGCACAGCGGGCAGGTACCCATCGCCAAGCGGGTGTTCAAGCTGCGGGTGGCGACCATCGCTCCACTGGCCTTCTACACCATCGCCACCCTCTCCAGCGTGGGAATCGCTCT
GGCCATCGCCTTCCTGGCGTTCAATCTGCACTTTCGGAAGCTGAAGGCAATTAAACTTTCCAGCCCGAAGCTGAGCAACATCACCGCAGTGGGCTGCATCTTTGTGTACGCC
ACCGTCATCCTTTTGGGCTTGGACCACTCGACGCTGCCCTCGGCGGAGGACTCTTTCGCAACGGTCTGCACGGCCCGCGTCTATCTGCTCTCCGCCGGATTCTCGTTGGCCT
TTGGATCGATGTTTGCCAAGACCTACAGAGTGCATCGGATATTCACTCGTACCGGCAGCGTTTTCAAGGACAAGATGCTGCAGGACATTCAACTGATCTTGCTCGTCGGCGG
ATTGCTTCTGGTGGATGCGCTGCTCGTAACCCTTTGGGTGGTCACCGATCCAATGGAGCGCCATCTTCACAACCTGACGCTCGAGATCAGTGCGACTGATAGAAGTGTCGTT
TACCAGCCTCAGGTTGAAGTTTGCCGTTCGCAGCACACGCAAACGTGGTTGAGTGTCCTGTACGCCTACAAAGGCCTTCTTCTTGTGGTGGGTGTCTATATGGCCTGGGAGA
CGCGCCACGTAAAAATACCTGCTCTCAATGACTCGCAGTACATCGGAGTGTCTGTATACAGTGTGGTCATCACCAGCGCCATCGTCGTGGTGCTGGCCAACTTGATTTCGGA
GCGAGTCACCCTGGCCTTCATCACAATCACAGCTCTGATTTTAACCAGCACCACTGCAACCCTTTGTCTGCTTTTCATCCCAAAACTCCATGATATTTGGGCAAGAAACGAT
ATTATCGATCCGGTTATCCACAGTATGGGCCTTAAGATGGAGTGCAACACACGCCGATTCGTGGTCGATGATCGCCGAGAACTGCAGTATCGAGTGGAGGTGCAAAACAGGG
TCTATAAGAAGGAAATCCAGGCTCTGGACGCCGAGATTCGAAAGCTGGAGAGGCTACTCGAGTCGGGACTAACCACCACCTCCACCACAACTTCGTCGTCCACATCACTCTT
AACTGGGGGAGGTCATCTAAAGCCAGAACTGACGGTAACCAGTGGCATCTCGCAGACTCCGGCTGCAAGTAAAAACAGAACTCCAAGTATCTCGGGAATACTGCCCAATCTC
CTGCTTTCCGTGCTGCCTCCTGTGATTCCACGGGCCAGTTGGCCGTCAGCAGAGTACATGCAGATCCCGATGAGGCGTTCTGTTACCTTTGCCTCCCAGCCCCAATTAGAGG
AGGCCTGCCTGCCTGCACAGGACTTGATTAACCTCCGTTTAGCCCACCAGCAGGCCACGGAGGCTAAGACGGGCTTGATAAACCGATTACGAGGGATATTTTCTCGCACCAC
TTCGAGCAACAAGGGATCCACCGCCAGCTTGGCGGACCAAAAGGGTCTGAAGGCGGCCTTTAAATCGCACATGGGACTGTTCACCCGCCTGATTCCCTCCTCTCAAACGGCG
TCCTGCAATGCCATATACAATAATCCAAATCAGGATTCCATTCCCTCAGAGGCGTCCTCCCACCCGAATGGTAACCACCTAAAGCCCATCCATAGGGGTTCATTGACCAAAA
GCGGTACTCACCTGGATCACCTTACCAAGGATCCGAATTTCCTGCCTATCCCCACTATTTCTGGCGGTGAACAGGGGGACCAAACGTTGGGTGGAAAGTATGTGAAACTGCT
GGAGACCAAGGTGAACTTCCAATTGCCCAGCAACCGGAGACCTTCGGTGGTGCAGCAGCCACCCAGTTTAAGGGAAAGGGTAAGGGGTTCGCCACGCTTTCCACACCGCATC
CTGCCGCCCACTTGCAGTCTCAGCGCCCTGGCCGAATCCGAGGACCGTCCCGGAGATAGCACCTCTATCTTGGGCAGCTGCAAGTCCATACCTCGCATTTCGCTGCAGCAGG
CCACCAGTGGAGGCACCTGGAAATCGATGGAAACAGCGGGCAAGTCGAGGCTTTCCCTCGGCGATTCCCAGGAAGAGGAGCAGCAGGCGCCTGCGAATGGCACCGAATAA
(SEQ ID NO: 14)

Start ATG: 1 (Reverse strand: CAT)

MRIIQPVQGTRYGPWPAVGLRLVLALAWATSAAAAMESSAELQALGYEAIRPGAASISTSSPSSSPPGESASTVTAGGTPIPPRSDWKYKRTKVKRRQQRLNSHSNLPGSTN
ASHAHHLLNLPPRQRYLKVNQVFESERRMSPAEMQRNHGKIVLLGLFELSTSRGPRPDGLSELGAATMAVEHINRKRLLPGYTLELVTNDTQCDPGVGVDRFFHAIYTQPST
RMVMLLGSACSEVTESLAKVVPYWNIVQVSFGSTSPALSDRREFPYFYRTVAPDSSHNPARIAFIRKFGWGTVTTFSQNEEVHSLAVNNLVTELEAANISCAATITFAATDF
KEQLLLLRETDTRIIIGSFSQELAPQILCEAYRLRMFGADYAWILHESMGAPWWPDQRTACSNHELQLAVENLIVVSTHNSIVGNNVSYSGLNNHMFNSQLRKQSAQFHGQD
GFGSGYGSRISIAATQSDSRRRRRRGVGGTSGGHLFPEAISQYAPQTYDAVWAIALALRAAEEHWRRNEEQSKLDGFDYTRSDMAWEFLQQMGKLHFLGVSGPVSFSGPDRV
GTTAFYQIQRGLLEPVALYYPATDALDFRCPRCRPVKWHSGQVPIAKRVFKLRVATIAPLAFYTIATLSSVGIALAIAFLAFNLHFRKLKAIKLSSPKLSNITAVGCIFVYA
TVILLGLDHSTLPSAEDSFATVCTARVYLLSAGFSLAFGSMFAKTYRVHRIFTRTGSVFKDKMLQDIQLILLVGGLLLVDALLVTLWVVTDPMERHLHNLTLEISATDRSVV
YQPQVEVCRSQHTQTWLSVLYAYKGLLLVVGVYMAWETRHVKIPALNDSQYIGVSVYSVVITSAIVVVLANLISERVTLAFITITALILTSTTATLCLLFIPKLHDIWARND
IIDPVIHSMGLKMECNTRRFVVDDRRELQYRVEVQNRVYKKEIQALDAEIRKLERLLESGLTTTSTTTSSSTSLLTGGGHLKPELTVTSGISQTPAASKNRTPSISGILPNL
LLSVLPPVIPRASWPSAEYMQIPMRRSVTFASQPQLEEACLPAQDLINLRLAHQQATEAKTGLINRLRGIFSRTTSSNKGSTASLADQKGLKAAFKSHMGLFTRLIPSSQTA
SCNAIYNNPNQDSIPSEASSHPNGNHLKPIHRGSLTKSGTHLDHLTKDPNFLPIPTISGGEQGDQTLGGKYVKLLETKVNFQLPSNRRPSVVQQPPSLRERVRGSPRFPHRI
LPPTCSLSALAESEDRPGDSTSILGSCKSIPRISLQQATSGGTWKSMETAGKSRLSLGDSQEEEQQAPANGTE*
(SEQ ID NO: 15)

Name: GABA B-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000012808751
TTTTCTGCCCGTTTGTTTGTGCTCCCGTCCAGGTGAAGGTTAAGTGCTCCGAGCTGAGGATCAAGGACAAGAAGGAACTCACCAAGCAATTGGATGAGCTCAAGAATGAGTT
GCTCAGCCTGCGCGTGGCCAAGGTGACCGGCGGAGCTCCCTCCAAGCTCTCCAAGATGTGAGTATGCCGCGGCACAACGACAAGCGTCACGAACTAATTCCGGTGTTTTGTG
CATTTATGTGCGTTCTTATCCATACAGCCGCGTTGTCCGCAAGGCCATCGCTCGCGTCTACATTGTGATGCACCAGAAGCAGAAGGAGAATCTGCGCAAGGTCTTCAAGAAC
AAGAAGTACAAGCCCCTGGATCTGCGCAAGAAGAAGACCCGCGCTATCCGCAAGGCCCTGTCTCCGCGCGACGCCAACCGCAAGACCCTCAAGGAGATCCGCAAGCGCTCCG
TCTTCCCCCAGAGGAAGTTCGCCGTCAAGGCCTAGAATGCATCCAGATACTCCGACTGTCTATAGTGCGCGTTAAGGGTTCTTCGGTACAATGAAGAAGAAAAACTAAATAA
AATCTTTTTTTAAAGTGAAAATAGCGATGCGTGTTTGATTTTGTGACCGAAAAAACGCGAGTGGCACATGGCTTTGGCCATGGGTTCATGGAAATTAGTTAAAAAACAAAA
AAACAAGTGCCGCATTCCCTGTTCTACATATGTTAGTCGGATCCAATGGATCTACTTAACTACTTGTGGTCCTTAAGATCATTATAGAGCAAAAACGTTGAGTTAGCGCCT
TGTAGAAGTAATGGCAAATCTTAAGTTCAATCTTTAATTCAGCCTTCTTTGAATATGCAAACTTGAAAAACACCAGAATTCTCAATTGGTCTACGTATATGTATTTAATTCG
ATTGTCTTTGATGTAAGCACGGCCATACATCGTAAATAATTATGAACTAACAAGTAATAAGTAGGTAATTGATATGAATTACAAGCTCTCGACTGCGAGCTCGATCATACTG
AATATGTTTGGCACATTTGGTTAATACTATTCTTCTGCACAGCGTGAAACAGCGGCGTGCTGCGGATGACGCCCGACAGCTCCTTGGAGTTCTTGCTTTGGTGATAGTTTCT
ACTCGGCATCGGCGACAACGGTTCGCCCCAGAACTTGAGCAGCGCAAAGATCTTGTAGTAGGCCACTCTGAAATAGAAGATAGATTCGCTATTAGCAGCTAAACCCGATAGC
AGCCAAAACTGCCAGCAAACTTACCTGTAGTTGCGATTGCTGGCCGCATAGATGATCGGGTTAATGACACTGGAAGCCCAGGCCATCACGGAGGCGATGATGTGCAGCCAGG
GGTACGAGGTATTGCGCTCGTCGTCCACCACATTGGCCAGCATTAGCGGCAGGAAGCAGACGAGGAAGCACAGAAAGATGGTGACCATCATGACGGTCAGTCGATTGTCTTC
GCGCGCCTTGCGCGTGCATGTGGTGGTCATGTAGGATCCGCCGCCGGACGACGAGGAGCCCTTGGCCGCTGCGATCTGGAAGTTATCGTGGTTGCGAATCTTCTTTTTCTGG
TGCAGCACCGTAATGTAGATGCACGAGTACGAAACGATGATCACCAGGCAAGGCAGCAGGAAGCCGATCACGAACAGCGTCTTCTTGATCGATCGCCCCTCCTTCTTGAGTA
TTGTGCAGGAGAAGGTGGCCTCGTCCAGGCCCATCTCGCCCCAGATGCCCAATATGGGCGGCAGCTGGGAAGATAAAAACCGATTAGTTGAATGTGTGCACACAAATCGAAG
ACGCCTTACTTACCAACAGCAGAAAGGAGACGGCCCAAACGAACAACAGCTGCAGGGTTATGAACTTAGGCTTATATATCTGCGAGTAGCGGCTGTGGCAAGCGATGAGTAT
ATATCTGCCAGGGGATTAGGGTGACGGTTGCGGATTAGCTTGCGACAATGCGGACTAGACTTGGGCCTACCTGTTCAGGGTGATGCCCACCATGCTGAGGAGTGAAACAGCC
ACATTGCCATAGAAGATCACCGGAAAGATCTTGCACAATGTGGTACCAAAAGTCCAGCTCTGTCGTTGGAATCGAGCAAATTATAATCATTCACTAATTGTATATTACATTC
ACAGGTATAATGACTCACCTCCTGGAAGAATCGCACTGCGGTCAGTGGCAGGCTGAAGGAGCAGAAGAGCAGGTCGGAGATGCTTAGCGAAATGACGAAGGCGGTTGTGGCA
TGCTCCCGTATCGTGGGGCTCTTGAGCAGCGCCAGCAGGGTAATCAGGTTGCCTGTGCGGCAATTCAGCATAGAGAAACAGGTTAGCTAATCGAGGAGCTGAATGTGGGATG
CCAATTACTCACCAAGAACGCCGATCGTCACAAAGACACAGGCACTAATGGCCGCGAATAATGTCGCCGAATGCGGGTAGATCGATTGCGTGGCGGCCGCCGGTTCATCCAT
CTGCATGTCTGCGTCCTGAAAGTAGCCCGTTGCCATGCCCATGTCCTGATCCATTCTGCGGGGGAAACAAGCAAACGATGTGATTAGCCAATTTCAATTACTTATTTTTGAA
CAGGAAAAGGAGCTGTATCTTTTAAATTGGTTGCATTCAAATCGCTTATGCTAATCCTATTGTTTTTTTTTTTTTAATTTGTTAGCCCGTATCATGAGTAATTAAAAAGGAA
GTTAATAGCCGGATGTAGACAACTCAATATGCATCCGCTTGGTTATTTGAAAACAATTAACTTATGAAGAGAAATGAATGAAATTACCATTCCTAATTGCCCAGAAGCATGA
TGCTTTTGTTAAGAAGGGATTGTTAAGAATTTAACTATCTAAACGAACTAATGTGTACATTTCTGTGCTCCGCGATAAAGGAAATGTAATCAGCAAAATCAAGTTTCAGGCG
CGTGGAAAATCTGTGAATGAATGAACGTACGAGTAATACAAAAAGATAACTAATCACCGCTTTGCGGATGCAATAGCCCAGTTTTGTCGTCGGTGACACCGAGTAAAAACCC
GATCCGATCGACGGGATTAGCGAGGAAGTCTGCGTACTCCCACCAGCTTATCGCATCCGAGACATTTGATTTATGTAAACGGATTTTGTTATAAGATTTCTGCCGAGCCAAG
CTGGCCAACTGGCGAATCGTAATGATCGAGCCGGTTTGCACGCCATTTGAATTCACATTTGTACGAATTTTTTTTTATCAAAAGTTCGAGTTTTTCACCAATTTCCTCATC

AACCGAGCAAGGCAAACGGCTTTGAATAATATGGTGTTATATATACATATATCAAATCGCTGCTGACTGCGTGATTGATGGCCCCAAGATTACATATTATCGAATCAGGATT
CAGAAGGAGATCAATGTCAAATGCGGACAGGAACATGAGAGACGCCTGTTATGCGCAATTAAAAATTTGGGTTTAATTGCTGTGGAAACTGTTGTTGGTGGCATCTTAAGTT
CCTGTTTAACAACATCAACTACTTATGTACGTAGAAGCGTTTAAGCCATTTGCATACAGATGAGAACTGGCTTTTGTGCTAATCAGCCAAGATGACTCCGATGATGATGACT
CATTACCTGACCAGTTTTCGCTGCTTTCTTTTCAACAACTACTTGTATATGTATTGTATCCAATAGCAATACATTGGATTTCCATGGTCTAGTCACGTATTATCATTTAATT
GACGCCAAGTCGTGTTATTGTTGAGCTATCGAGTTCAGCTCAAACATTTCTTATTCCCATGAATAAGCCGGCAAAAATATGCAATCTATGAAAGTTAATATAAGCAAACCTT
ACTTTGACTCAATACCAATGCACTTTGTGTCGATAGGTTCACGCAATTGAGGCGATTATTCCGATAACCCAAGCGATTGACTGTTCCCGTTTCGATTCCAATTGAAATTTGG
AAATGTACAATAGTTTTGCTATATGCTGTCAAGTACGCTCTTATCTTCTCTGGGTTTCTTCAGAGTTTCGAAACGCTTCTTCTTTTTTTTGTTTTTTTTTTTTTGGAATCT
CGTATTTTGGAAGGGGCTCCCCTCTGGAATTTGTTACACTGTCGTTATCATTGCGAACAAGCGGCCCGAAGCTATCAGCGACTTTAACATTTACAATGCACTTTTTTACGAC
CAATTAAATGTACATTTTCCTTTCTTCGCCCGTTGATAAGCGAACGCGATGTGGCGCAGGCAATGTGTTGCTCTTGCGACACAAACGCAATCAAAATGGATTCAATTTCGCT
TTTTCCCAGTGAAACGAAGAACGAACCGACCATCATGATATGCTGCTCTGCATGTTGCGTATTGAATCAATGACAATTTCAATTAAGCCGCCCGTTCGTCATGCGTTTTCGT
GGGCTTCGAAATGCTGATAACGCTGCTGTCCTCCAACTGCTTTGCATGTGGACACAATTCCATTTATTTAATTCTTTTATTTGGATCGGTTAAATTAAAAAGCGCCTTGTTA
CGCATTTAACGTTGTTTCCGGTGCGTGGTGGTTTCATGCTTCTGGGAACGGCAAATGGGTTTAGGATTGGGAACCCCTCATCATCCGTCCAAATATGGGGAGTTCAAGGAAT
GGAGGGGGGAGGGGGGGAGGGGGCGGAGCAGCGGTCCGCTTTCATGCATGATCAACATGTGTGTTTGTTTTCCAACTCTATGTGAGTTGTTTCAGTTCTTTTCGGTCTCCTC
GGTTCTACTCAGTTCGTTCTTACTTTCTGGCTTTTGTTTCCGAATAGGGGGAGCGCATCGGGCCGCAACTGGAAATGGAAATCTGAATATGCATCCCGGAGCATCCCGTCCG
AAATGAGCACCATGCGCCATGCCGTTTCGTTTAATTCCCTACCCATAAATGCGTGTCCAAACAGCAGAGGGCATTTCAAGAGGTCCAACTATTAGCTATGTGCATATATAAT
TCTCTAGATATATTAATTCAGCTAAAGTGCTAACCGCGAGTGTGCCGCCAGTTGCCCCTCAACAGTTTGCAACGATGATGATGGCACAGATATCAGAGATAACACTCAGATC
GCGCCCCTCACAAGTATGCTCATTATGTACATTCATATGTACATATGTACACACAGCTAAGCACATAAGGAATACGATATGCGCTATGTACATATATACATATTTACATCAA
AGTCTTATCAGCCTCGGGTTCCCACCGTCCAAAAAAAAAGGCGGAACGAAGCTTTATAAATAATCAATCAACAAATATTTACAATCAAGTATTTATGTAGCGGGCGGAAAGT
CGAATGCAAATAAGGTTAGACGCATTTGTACATATGTATTTTCTTAAAGGGTTCCTCTGATCAATATCTGCCCGTTTTTATCTAGCCCCGTTTATAAGGCTGTGAATGCCCC
GATTGTGTTGGTCGTGATATCAATTAAATCGGCATAATCTCATGTCATCTTTTTAAACATTATGATAGATAAAATGTATTATAACGACTCTAATTACATTGGGATTAAATGG
GTTTTCGGAAAATACCTATGCAGTAAATCAAATTTAAGATCAATGTGCAGTTATTTGGGGTTTCTTGAGTTAGGACCTCTGTTTTAATATTGAGAATTACTCGATATTATCT
TCCGTTTTATTTTATTTTCTGATCGTGTCGGTGCACGAGAAATACATATTTCCCACTTTTGTTGCATTTCGCTGACTAATCGAAAGCTAAATACCAAATAAGTTTCCGATAT
GGGGGGTATGTATGTACATACATACATATATAGTACGAGTATATATAATATACACTGAATAAGCCCGAGCATACATATGTACATATGTAGATAGTGCACATGTGTATTGTGG
ATCGTTTTGGAGCTGGGTTTTTCTTTGCTTTTCGTCGTTTTTAAGAGACTTGAGCTGAATGAGTCGCTACCAACTGTTGCGGTGGGTTAAAGGGTGGGGGGAGGGGGCGAAT
GGGGATGGGGGATGAAGATGATGATTCATGCAGCTGATGTTTGGACAGGGAAGAAAGAAAGAGCGAAAGAGAAAGCAAATCGGAAGCGGCGGTATCAAACAATCTGTTTAAA
TAACAACAAAAGACTTTATAAGCCTGACTTTCATTTCCATTTCCAATTTATTTTTCCCAAGTCTGGCTCATTTTGTGCAAATAAACAAATGACTCATTTGTTTAAAAATTCG
ATGAGACTCAAGGAATGTCACATTCACAAAAATTTTCAATTTGAAATCTGAAAAGATTCAACAGCATAAGCAAATGAATTTAAGATTGCGATATTTGGGTATTATCCTTAAT
TGGCTGGCTGCCTAATGTAGATTTGTTTTTCTCGGTGTACCGATGCATCCTCTTTGCCCATCCACTTCTTTCTCCCCTCTTTGCTGATTTTTTTTTTTTTTTGTTTTGTGAGA
GGGTGGAATTCGGGATTTGGGATGTGAGCGAGAAGGATGGCCATATTTTGAATTGTGGAAACAGCCGTTCCATTATCGTTCCATTACAAATGCACACACACACACACACACA
CGCGCGCGTGCAGTTTGCGCTTCTCTGTGTGTGCGTTTGGACCAAATGCAATAACAATAAAAACGCGCAACTCAAATTGTTAACAAAGGTAAGGCCGTTTTTTTGGCCGGAT
TTGGATTCTGGTCTTGAAAGAGAAACACACACACACACACACGCACGCAAAAAACCAACGAAAATAATAACAAAACAACAACTCACCCGAACTCGAACCCGTTTCACTGTTT
TCCGCTTATTATAAATTTATACTTCGTTTGGCGTTTCTTAAATGCTTTCGAATTGCATTTTGGATTTGACATTAAATCGCTCCGCGCAGCACAGTAACAAAAATAAACGGAA
CAAAATAAATACAACCAGCCGATTCTGGCCTTCTCACAATTCCGACTTTTGAATTATCGAAACAGAAGCGCTGCTTGGGCCAACTAAGCCAACTTAACTAACCGATCGCAGA
GCGCAAAGTGTCGTATTTTCACTGTCGTCAATTAACAACAGCGGCAACAAGAACAACAAACGAAATGAGCGGCTGGGCTACAGTTTTTAGTCCGAGCGCGATCGAACATGT
TGGCTGCCTGGCAACAAGTTCGATGCGAACTTTTCTATTGATCTAAGAAATCGGCGTTCTATCGATTGACGCGCTGCGCAGGCGCCTATCGATTTTTGAAATTCTAAGCCCG
CCTTAAAGGATATTTGTAGAAAAGTAGCTTGCAAAGAAATTAAATTCCTTGAAATAAAAAAGAAATTTAGTACTCTCGTTCTATACTCACTTCTATTCACTTGCGGTTAGTT
AGTTAGTTTTTAATGCCACATTTTTATGATGATGACTTTTTTTCAGCTCATCTATGCGCAATCAATTTTGCGCCAAGCGACAGCTGCACATAGTTTCTTTTGACCAAGCAAT
CATTTATCATATAATTTTCCAAGTAGTAGTGCCAAGTTCATAACTGATGGAGTGTGCTGCCAATTTGCATGTTGTGGAGCCATTCAGTTTCAAGGTCGCGCTAATTAAATAC
TCGTTACAAGTGATGCCAAATGTATAAATATTTGAACAGAACGTATTAAATTAATATTATCCACTAAGAGCACTATTTTAAATTCTTAAAAACCAGATGCACGTCAAATAGA
TGTACATAAAAAGTATATTTACTTACAAATATCATACACGGACCTGTTTTATTCCTCATCACTGGCCATTAGTTGATAACATGTGGACTTCTATGATATCCTGCTGCATTTT
TGCCGTTTAGTCAGGATTTAGAGCTTGTACACAATGCGACAGCTGAAAGGGCGAAATCGGTGCAATCGTGCTGTTCGGCATCTGAAAGTTCAGGGCAAGATGTGGTTGAAAA
ATCTTAAAAGCGGTTTGGAACAAATTCGAGGTGAATCTATCCAAATCGTTAGAATCCCATGAGTACGTTCATC
(SEQ ID NO: 16)

Exon: 6913..6695
Exon: 2519..2365
Exon: 2292..2147
Exon: 2075..1975
Exon: 1908..1806
Exon: 1744..1257
Exon: 1187..1001
Start ATG: 2500 (Reverse strand: CAT)

Transcript No. : CT10621
GGCCCAAGCAGCGCTTCTGTTTCGATAATTCAAAAGTCGGAATTGTGAGAAGGCCAGAATCGGCTGGTTGTATTTATTTGTTCCGTTTATTTTTGTTACTGTGCTGCGCGG
AGCGATTTAATGTCAAATCCAAAATGCAATTCGAAAGCATTTAAGAAACGCCAAACGAAGTATAAATTTATAATAAGCGGAAAACAGTGAAACGGGTTCGAGTTCGGAATGG
ATCAGGACATGGGCATGGCAACGGGCTACTTTCAGGACGCAGACATGCAGATGGATGAACCGGCGGCCGCCACGCAATCGATCTACCCGCATTCGGCGACATTATTCGCGGC
CATTAGTGCCTGTGTCTTTGTGACGATCGGCGTTCTTGGCAACCTGATTACCCTGCTGGCGCTGCTCAAGAGCCCCACGATACGGGAGCATGCCACAACCGCCTTCGTCATT
TCGCTAAGCATCTCCGACCTGCTCTTCTGCTCCTTCAGCCTGCCACTGACCGCAGTGCGATTCTTCCAGGAGAGCTGGACTTTTGGTACCACATTGTGCAAGATCTTTCCGG
TGATCTTCTATGGCAATGTGGCTGTTTCACTCCTCAGCATGGTGGGCATCACCCTGAACAGATATATACTCATCGCTTGCCACAGCCGCTACTCGCAGATATATAAGCCTAA
GTTCATAACCCTGCAGCTGTTGTTCGTTTGGGCCGTCTCCTTTCTGCTGTTGCTGCCGCCCATATTGGGCATCTGGGGCGAGATGGGCCTGGACGAGGCCACCTTCTCCTGC
ACAATACTCAAGAAGGAGGGGCGATCGATCAAGAAGACGCTGTTCGTGATCGGCTTCCTGCTGCCTTGCCTGGTGATCATCGTTTCGTACTCGTGCATCTACATTACGGTGC
TGCACCAGAAAAAGAAGATTCGCAACCACGATAACTTCCAGATCGCAGCGGCCAAGGGCTCCTCGTCGTCCGGCGGCGGATCCTACATGACCACCACATGCACGCGCAAGGC
GCGCGAAGACAATCGACTGACCGTCATGATGGTCACCATCTTTCTGTGCTTCCTCGTCTGCTTCCTGCCGCTAATGCTGGCCAATGTGGTGGACGACGAGCGCAATACCTCG
TACCCCTGGCTGCACATCATCGCCTCCGTGATGGCCTGGGCTTCCAGTGTCATTAACCCGATCATCTATGCGGCCAGCAATCGCAACTACAGAGTGGCCTACTACAAGATCT
TTGCGCTGCTCAAGTTCTGGGGCGAACCGTTGTCGCCGATGCCGAGTAGAAACTATCACCAAAGCAAGAACTCCAAGGAGCTGTCGGGCGTCATCCGCAGCACGCCGCTGTT
TCACGCTGTGCAGAAGAATAGTATTAACCAAATGTGCCAAACATATTCAGTATGA
(SEQ ID NO: 17)

Start ATG: 239 (Reverse strand: CAT)

MATGYFQDADMQMDEPAAATQSIYPHSATLFAAISACVFVTIGVLGNLITLLALLKSPTIREHATTAFVISLSISDLLFCSFSLPLTAVRFFQESWTFGTTLCKIFPVIFYG
NVAVSLLSMVGITLNRYILIACHSRYSQIYKPKFITLQLLFVWAVSFLLLLPPILGIWGEMGLDEATFSCTILKKEGRSIKKTLFVIGFLLPCLVIIVSYSCIYITVLHQKK
KIRNHDNFQIAAAKGSSSSGGGSYMTTTCTRKAREDNRLTVMMVTIFLCFLVCFLPLMLANVVDDERNTSYPWLHIIASVMAWASSVINPIIYAASNRNYRVAYYKIFALLK
FWGEPLSPMPSRNYHQSKNSKELSGVIRSTPLFHAVQKNSINQMCQTYSV*
(SEQ ID NO: 18)

FIGURE SHEET 11

Name: Melatonin receptor-like 2
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384305
GTTCGTTTGATCGGTTGACCTCGATTTATTTACTGCTATGCAAAGCTAAAATTCAAAAATCCCTACTCATAGATGAAGTTTTTTGCACACTTTAGGTGTAACTTGGTGCTTT
AACTTGGCGTTTCTAGTCTTAAATAATTGGCAAGATTATCTGCCTAACTTTTGGTGACCCCAACTTTTAATCACACACCAAGCGGAATTATACTAAGGTTTATTTCCCTTGG
CATCAGACTGTCTTCTTCTTCGCTCTTCCAGCTCTTGTCTTTATTAAACCTGGCCAAGGACCTGTAACTTCGTTTAATTAAGTAATTCAATGAAATTGACAGCTGTTTGCTT
GGGTCAAGCATAATTGATGACTACTCATAATCAAAACAGAAGTGCTGCCAATGAATGGGAAGCGGAAGTCTGGCATAAATTACAGGTCTGTCCATCCGAAGCGCAATCATCA
TATAATTGTCATGGAACCACACTTGGCACTCACTTTTGCCTGGCTTCTGCTCCTGCGAACGCTGATGTTGCTTTGTGCTTTGTGCTTTCTGCTTGATGCTTTTTGCTTTGTG
CCCAAGTCTCGTGACTCCACTCCACTCGGAGTGGCATCCATGGTTAGGAGAACAGCAATAATGAGGCACACCAACCGGCCAATTGGCAATTGTGCGAGGTCAAACACAACCG
GTGGAGGAAACTGGTGTCGGCCGCACAGGGAAGACTTAAATTTTCTTTAATTGCTCGAATTTATGGGTTGGTAGTGGCATAACGACAATTGTCCAAGTCAAGGTCACAAATG
ATGCTCCTGCCTTTTTAGTTCGATACACACAAAGAAAAGGGGACCATACATATACAAGCAACCAACTAGATGAAACCTATTTTAAGGACGGTTGAATTGTTGGCTTTTAACA
TAGTAGATTAACTTCAAAATTAGCAGTTACTACTACTAATAACGAGCTAAGCCACTCATTTGCGTAAACACTTTCTATCTGTGCACTTTTATCTGCTTGAATAGTGTCTGGG
TGATGGCCTCGATTGCCGTGGTTGGCAACCTATTGGTCCTGCTCGGTCGCTACTTCTACAAATCACGGAGCAACGTGGAGCACTCGCTCTACCTTCGCCATTTGGCCGCCAG
TGATTTCCTCATGGGCATTTACCTCACATTGATTGCCTGTGCGGATATCAGTTTTCGCGGCGAGTATATCAAATACGAGGAGACCTGGCGGCATAGCGGCGTTTGTGCCTTC
GCAGGTGCGTATCCTTCGACATCCTAGCCATTTCCATTTCGTAATGGAATGCTAATGTGTATTACGTATTCGCCGTGTAGGCTTCCTCAGCACCTTCAGCTGCCAGTCGTCG
ACGCTGCTGCTCACATTGGTCACCTGGGATCGTCTGATGTCGGTGACGAGGCCCCTCAAGCCACGGGATACGGAAAAAGTTCGGTAAGTGTGAACTACAATAGATATGTAAT
AAAGTGCTGGGCAAAGTTATTGTTTATTCTATTACATTAATTACTTTTGCTGCGGACATTTGCTGATTTGTTGACTGCTCAAATATCAATGCACAGAATTTGTTTACCCTAC
TTGAATTTTGTTCGTAAATGCAATTATGTTTGTTTTGACCGCCGCCTTACTATGTACAATATATATTTTCTGTTGCTTCATTTTAGTTTGCAAAAGCTATGGTTTTGTATGT
GTAAGCTATGATAAATTATTGAAATGACAACTTCCACAGCATTGTCCTACGCCTTCTGCTGTTGTGGGGCATAAGTTTCGGATTAGCTGCAGCTCCACTTCTTCCCAATCCG
TACTTTGGTAGCCATTTCTACGGCAACAATGGTGTCTGCTTATCGCTGCATATCCACGATCCCTATGCGAAGGTAATACACCATTTCCTTGGAAACGCCTGTTCTAATATTC
CAAAATGTATATTTTTAGGGTTGGGAGTACTCGGCGCTGCTGTTCATCCTGGTCAATACGTTGTCACTGATCTTCATCCTCTTTTCCTACATTCGAATGTTGCAAGCGATAA
GGGATTCGGGTGGCGGAATGCGAAGCACTCACAGCGGTCGCGAGAATGTGGTAGCAACTCGGTAAGTTTGGATAATTCAAGTAAAATGATTTAATTATTGAAATATCATGTT
TTTACAGCTTTGCCATCATTGTGACCACCGATTGCGCCTGCTGGCTGCCCATAATTGTGGTCAAATTGGCTGCTCTTTCAGGCAAGTTCATCGGCAACGCCTTTAAATGATT
TATTTTGATTATCAACTTATTCGTATATATATAACCCCTTTCAGGCTGCGAAATCTCGCCCGATCTTTACGCCTGGCTGGCGGTTCTTGTGCTGCCAGTGAACTCGGCCCTC
AATCCAGTGCTCTACACATTGACCACTGCGGCCTTTAAGCAACAGCTGCGTCGCTACTGTCACACCCTGCCCAGCTGCTCGCTGGTGAACAACGAGACCCGATCCCAGACCC
AGACTGCCTACGAGTCCGGACTGAGTGTCAGTCTGGCGCACTTGGGCGGTGGTGTGGGCGGCGGGTCGGGGCGAAAGCGGATGTCACACCGGCAGATGAGCTATCTGTAGGC
GGGCGCAACTGGCTTAAAATGCAGCACGAGATCAACAGACTGGGGCTAGTGGGTGGGTGGCATTACAACAGGACACTTCCATGTAATTGGTCCCTTAGACTATTGGTATTAT
ACGTACTTATTGATACTTATTCTAATATCAATTTTAAATTACTTATAACATTTGAGGATATGATAACATAAACCTTAGACAAGAGATTCCAGCTATTTAAGTTTCTTTTTTG
TATATTATTTATTTTACCAATCGTTTTAAGGCAAACCTCTTACAACTGTCCAAAGAATATCCTAACTATTTGTTGCATCCCCATTTATTCGACCTGCCCAATGGGATATTGT
TACATGTGATCTAATAGTTAAATCAAATTTTAATTTTATAAATACATATTCAGATGTTAATTGAAGCTTTTGAACATAGGACTTGCAATCGAAGGCCGCAAATGCCTGCCAT
CCGGCTTTCTATGCGTTGCCCTGAGGTTCAGTGGCTACATCGAGATCCGCCTGGACGTGATAGGGTCCCTCGTTCACATAGGCCGTGAAACGGATGGTTTCGGTGGGCGAAA
ACTCAGCGCTAGCGGACAGAATCTCCATAATTTGCCACTTCATATACTTAATGTCCCTGTGCAGATCCTTTTGACGCGAAAGAAGGAACTCCACAGCACCGCCGGTATTCTT
TAGAGCTACAATGCCCCAGTTGTACTTACAAATGGTTTTCAAGAACGACAAGGCGGCCATTTGAAGATCGGGGAACGGAGTGTTTATCAGGTCCATAATGATATTAGCATGC
GCTCCCCCAGCGAAGCACTCGTACCAGTTTTTACCAATGTTACTGAAAAAGAAGAGAATCATTAGAAGAAAAACTTCAATTATAGAGTTAAGCATTACATGATTTCTGTGGC
AGGCGGCGTCTTGAAGTCATAAATCACATCCAGGGAGTTGAGCAGCCTCTTCTTGATGTGCGCCGATAGCTTCTTAGTATGCGAGCCATACTTCTTAAAGGA
(SEQ ID NO: 19)

Exon: 1001..1236
Exon: 1313..1427
Exon: 1720..1864
Exon: 1923..2077
Exon: 2136..2209
Exon: 2285..2574
Start ATG: 1011

Transcript No. : CT13764
TGTCTGGGTGATGGCCTCGATTGCCGTGGTTGGCAACCTATTGGTCCTGCTCGGTCGCTACTTCTACAAATCACGGAGCAACGTGGAGCACTCGCTCTACCTTCGCCATTTG
GCCGCCAGTGATTTCCTCATGGGCATTTACCTCACATTGATTGCCTGTGCGGATATCAGTTTTCGCGGCGAGTATATCAAATACGAGGAGACCTGGCGGCATAGCGGCGTTT
GTGCCTTCGCAGGCTTCCTCAGCACCTTCAGCTGCCAGTCGTCGACGCTGCTGCTCACATTGGTCACCTGGGATCGTCTGATGTCGGTGACGAGGCCCCTCAAGCCACGGGA
TACGGAAAAAGTTCGCATTGTCCTACGCCTTCTGCTGTTGTGGGGCATAAGTTTCGGATTAGCTGCAGCTCCACTTCTTCCCAATCCGTACTTTGGTAGCCATTTCTACGGC
AACAATGGTGTCTGCTTATCGCTGCATATCCACGATCCCTATGCGAAGGGTTGGGAGTACTCGGCGCTGCTGTTCATCCTGGTCAATACGTTGTCACTGATCTTCATCCTCT
TTTCCTACATTCGAATGTTGCAAGCGATAAGGGATTCGGGTGGCGGAATGCGAAGCACTCACAGCGGTCGCGAGAATGTGGTAGCAACTCGCTTTGCCATCATTGTGACCAC
CGATTGCGCCTGCTGGCTGCCCATAATTGTGGTCAAATTGGCTGCTCTTTCAGGCTGCGAAATCTCGCCCGATCTTTACGCCTGGCTGGCGGTTCTTGTGCTGCCAGTGAAC
TCGGCCCTCAATCCAGTGCTCTACACATTGACCACTGCGGCCTTTAAGCAACAGCTGCGTCGCTACTGTCACACCCTGCCCAGCTGCTCGCTGGTGAACAACGAGACCCGAT
CCCAGACCCAGACTGCCTACGAGTCCGGACTGAGTGTCAGTCTGGCGCACTTGGGCGGTGGTGTGGGCGGCGGGTCGGGGCGAAAGCGGATGTCACACCGGCAGATGAGCTA
TCTGTAG
(SEQ ID NO: 20)

Start ATG: 11

MASIAVVGNLLVLLGRYFYKSRSNVEHSLYLRHLAASDFLMGIYLTLIACADISFRGEYIKYEETWRHSGVCAFAGFLSTFSCQSSTLLLTLVTWDRLMSVTRPLKPRDTEK
VRIVLRLLLLWGISFGLAAAPLLPNPYFGSHFYGNNGVCLSLHIHDPYAKGWEYSALLFILVNTLSLIFILFSYIRMLQAIRDSGGGMRSTHSGRENVVATRFAIIVTTDCA
CWLPIIVVKLAALSGCEISPDLYAWLAVLVLPVNSALNPVLYTLTTAAFKQQLRRYCHTLPSCSLVNNETRSQTQTAYESGLSVSLAHLGGGVGGGSGRKRMSHRQMSYL*
(SEQ ID NO: 21)

Name: TSH/FSH receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013383844
TTTTCCGCCAAATTCGCTGGCTTTTAAGTGCTCACATAAATGAAATTTTCAAATAGCAAATAACATAGGAATAATCCTTCTAAATCGCAATATAAAAACATATATACAAGGA
GTGGTAAATGACAGTATTCTATTAATGGCAAAACGATTTATTTAAACAGACGTCGCATCCGCATAAACGATCTTTCACACTTTTTATTAACGGTTAGTATAAATAAAAAAAA
ACATTAATCCTTAGTTTATTTAACCAATTTTGTTTGCTGAAATCGATTCTTGAAAGAGAAGCTGACATTTTTTTATATATGCCTCCCTACTTCAGGCAGTGCGAACATACC

GCATCAACGCGACTTGTAGCGGCCGTGTGAAAAGGCTATTTTGTGGTTAGCGCAGATCTGCGCTGCAGTGTGACCCCGGCTACAGGTATTTCGGCATCGCACAGGGTCACAC
TCGCCCAGGCACCCGTATAAATCCCCGCCCCCCTGATCGACGATAGTATTTCGAAAAAAACGCGAGGAAAAGTCTGATGCGCGTTAAAATCGGAGCGCAAATCGTGCATAT
ATTTTTTTTACACACCGCGCGCCGTGTTTACGACTGAAAACGGAGACTGACTGAAAAGCCTTGGCGCACAAGCACGCAATGAAAACGCAACGATAAGAATGTGATATTGTG
ATAAATACATGATTCGCTGACAACTCTATTTTATATTGGAATGCATTCGAAATCGAACGCAGCGCAGCGTGTGTTGTTTACTTAACGATTTTCGGTGTTTGTTTGCCCAACT
TGCGGCGGTGAAATTCAGCATTTTCTGATTTTCCATTTGCCAACGAAAGAGAGGAAAAGTGAAAAGTGCTATTCGACGAAGCGAAACGGAACGCAAAATCAAAAGAAAAACA
GCATAGCTAAAGCTATGTGTAGTACATAGCCACATAAATTGCAAATTGAGGCCAAGAGACAAGAGAACCTGAAATCCACACAGATCTTGACGAGTGAGCTAACGATGGACAG
GAGTCGGAGTAGCAGAGCTGCCAGCAGCAACCAGTTCATCAGGCCATGTGGCCTACTGACGACCGTCATCCTGCTGCAGACGCTTGTGTCGATGAGTTTGGCCATCGAGGAG
ATGTCGCCACCACCGGCCGCCCCGCCACGTCCATCGCCGCCGCCCACCGTGAAGCTCAACAAGTGCTGCCATTCGGGCGAGTATCTGAACGATGGCACCTGCATCGCCGGCA
GCGAAGCCCTGTGGCTGCCCATGGTGTACTTGGTGCAGCAGCAGCGCTTCTTCGAGCCCCATGGCGCCAGTCCGCGATTCTTGAAGTTCCTGCCCAATACGAGACCCACGTG
CCGAAAGGACCAGACGACGGAGATCTTTCGCAGCCGCGGAGCCAATGTGATGCTCTTTCCCAATGGCACGCTCTATGTGCGGGAACGCGCTCTAATGGTCCAGCCGTCGGAT
TACTCTGTCGACTGGGAGGTGGCCGTCGTTTGCTTGAACGACAGTCAGCCGATCAATGCTCTGGAGGATCCCGATTATGCAGCCAATCCTCTCGTGCAGCAGGAGCCGCCCA
AGGCTAGCCTGCGACTGAGCAAGTGCTGCGGCAAGTGGGGCAGCTATAACACGCAGCTCCAGAACTGCGATCTGCAGCCCAACCATCAGGCAGCCGTCGATGGACTGCTGCG
ATTGTCGCCGCAGCTACCGGAGGGCAGCTATCAGACTAGCTACGGCCTGCCCGACTGTGGCCAACCCGGTGGCTACTCCATTGCCGGTGATTGGCAGGATGCCAAGCTGGAT
CGGAACACGGCCATGCTCCAGTTGCCGCACAAGAACCTCAGCGCCGGACAATACTGCCTGGAGCACACGCAGCGCGAGGGCGAGGTGAAGATCATAGCCTGCCAGCATTTGT
TTAGCTCCGCGGCAGGAGCCGGCATCCACGACGGCTCCATTGGAGGCACCATCGAGCAGGCCAATGGCCAGAATCTGCAGAAGGCCGTCCTGACTGGTGGCATACTTGTGTC
CATTGTCTTCCTGTCCGCCACTTTGGTGGCCGGCTTCCTGCTGCCCGCTGTGCATCATGCGCTTCATTGGCGCTGCCAAATCTGCTATGTCACCTGCTTGCTATTCGGCAAA
ATACTGCTGGCCATCGAGGAGCTGAGCTCCAGCCTGCAGCCGGGCAGTGCCGCCTGCCATACGCTTGCTATCACCATGCAGTTCTTCTTTCTGGCCGCCTTCTTTTGGCTGA
ACACCATGTGCTTCAACATCTGGTGGACGTTCCGGGACTTTCGGCCCAGTTCGTTGGAACGCAATCAGGAGGCACTGCGTCGCTATCTGTACTCCCTGTACGCTTGGGGCGG
TCCTCTGCTTATCACCTTCGTGGCCGCCTGCGTGGACCAGCTGCCGGAGACGACGCTACTGCGTCCGGGATTCGGACAGCTCTACTGCTGGTTCGACAATCGCAATCTCTCG
ATCTTCGCCTACTTCTACGGACCCATTGGCCTGCTGCTGTGCGCCAATATAGCGCTCTTCGTGTCCACCACCCATCAGTTGACGTGCGGCCTGTGGAAACGGGACGATGTCA
AGTCCTCATCGGAGAAGTCCGCCCTGGGACGCGTCTGCCTGAAGCTAGTCGTCGTTATGGGTGTCACCTGGATAGCGGACATTCTGTCCTGGCTAGTGGGTGGACCGCATGG
CGTATGGTTCTTCACCGACCTGATCAACGCCCTGCAAGGCGTCTTCATCTTCATCGTGGTGGGCTGCCAGCCGCAGGTGTGGACCGCCTGCCGCAGGATCTTCTGTCCGCGA
TTGCGTCATGACATCACCAATACGACCAACGGTGTCCAGCATTCGAGCAGCTCGCAGGGTTTGCCCTCGATGGCCGGCGGCACGGAGATTACACAGAACACCACCACCACCA
CCACGACAACCAACACGACCGCCACTCACATGCCCAGCAATCCGGCGGAGGATGAGGTACCGGAGAAGGCTCCCATTGCCCCCGTCGCACCCATCGTCAAAATGGAGACCAT
TTGCTAAACGTAAGTTGAACACCCGATAGCCAATGAAGTGTGCTTGTAAAGAGATAGCTCTGTTTGCCAGCGCGCAAACAAAGCGCCTGCTGTTCCCACTCCCAGCGCAGTT
TGCCAAGTGGCTTACAAAACGCGGCTTTAAGCACACACACACACACACACACACTTGCTGATAGACTTATCGGCCTTAATGGGCTTTAGGGCCCAGCCGGGAATGCAAAGTA
TCTCAATGAGCTGGCTGGCTGGCTGGCTGGCTGTTTGCTCGCCGGTTACGCCATGCATCGCAGCAGTTTCGATCCTATAAAGCGCACGTATGTACAGGTGGGGCCACTTCAA
TATCACTAAGCAGCTTTAGCTTCAATAGACTTCAATCGGTGGAAGGTTTGGTATTTGGACCCAAGAACCTAAGTTTCTTATTAAGTTTTAGTCTTTCAAAAATAGCGTGCTA
AATAGAGGGTCTTAAATTAACTCGAAAATGATTTAGAATCCGTGTAAATTGAATTGACACATTAGCTGAATTTTCCGGTATATAACCAAAAATGTCCGGACGATATATGTGT
CGTAGAAGATAGTGTGATAATTGCGCTGTCTGTGGATTTAACCCAAATACCAATGAGGTATTCACAAAATGTGAAAACAAATAGGAAACCCCGAATTTGTCATGCGCCGACA
ACTTTCGAGAACCAACCACCACCATATGAAGATAAAAAACAAAAGAAACTAGTTTGGTTTTCTATGGGGGAAAACGATACGAACTTTTCAATTCAAAACCTTTGGAGTTTCA
ACGGTCAATGGCCATGTTGGAAAATACCCTTCTCTATAGGACACTATTCATCTCACACCCCCCATGTGGACTATACGCTATCCAGATGAAATTTAAGGGTTCGAGATTTCCT
GACAACTTGCGGCTACCTGGAGCGTTTTTCCAACCAGGCGCCGGTTGAAAAAGTCTAGCACGTTTTCGTTTCGTTTTATGGTTTTATAGGGGGAAATTTACGTATTTACGTA
GGATAAGAAGATCGTCCTGCGATGGTCAAATAGTGACCATATTCTTCATTTGCATGACCGAAATTGTTTAACAACATGCGCCAAGTTTACGATTTCGCGTGCACCACTTGAG
TTCGAGTATGGTCATAAAGATGAAAATGTTGCCGCTGTGTGCCGTTATGTTTTTCATCGGAAAACTCTGACGCCGGGCCGGGCGAATTCGATTTTTGAATTTCCTGTCGATT
TCCACATTTCCACATGGCCATAAAGTAATCAGCGCAACCCTTGACACACACGCAAGCGAAGATAATGGAAATCGTTGCTGGCATCCGTATCGGTTGGATTCAAAAAATGTGT
GTTGCCGAATCCGTAAGGGAAAATCGAGTTTAAATTGGACTCACCGGCGACGATGGCGATGGTGATGGTGATAAGCACACAACACACCATCGTCTTGTTTGGATTTCGGCCC
ATTTACTTATTTACTTGTTACTTTGTTAGTTTGTTAGTTTGTTAGTTTGGTCGCCTCGAGTGGTTTTGACACATCTCCCAAATGATCCAATGGCAAACATGCGGACAGGTCT
CGAACTCGGCACAATCTACAATTATCTAAAAAGAAACCTCTGTGCGGCGATCATAAAGTTTTGAAAAACAATAATTCCATTTTCGGGCACTAAAAAAAAAACAGGGCAACTT
GAAATCTAGCATTAAACTGGTGGCTTAATGTTCCAATGAAAAGATAAGGATTTAACAATAAAAAAAAACTTTTTTTTGTATGCACTGCATGTTGATAGCAACATTAATGAA
AGAAACTAATCAATGAAGTTTCTGTTTCTCTATATTAACAACAAATGCTGCTTTGTTTAGGCCATGTTTTTGTTATTAGAGCCCAAGAAACATTGGTTTGCATTGCACAAAG
GCGTATGGAAAATGTGTAATTACTTCAGTTTGTAACATTTTCGAATATGACAGCTGCTGGTGGCAGTTAATTCTTTGCCCTCGCGTTTTATTGTTTGTTACTTTTTACTTTT
ATTTTTATTTTTTTTTTTGGCCCTGATCGCGTCATACTTTTTGGGATGCGTGTCTGTCCGGGTGAACAGATGTGCTGTCTGTCTGCCATTCGCTTGTGTGTAGGTCCGTG
TTGGAATCAGAGGTTTTGCCCAAACGAAAGAGAAAACAAAACAAACAGTGGGTTTTTGGCCCATGGGCCCATTCGATGAGATTGCGCAACCGAGTTAAATGGCCTGGCCCCT
GAAAAGCTGAAAACCGAAAACAGAAACATGAAATCTGAAATCTGAAATTGTGGAAATCTGAAATTGTGGCTGGCACACGCAACACTCGTAATCAGACTAAAAAAAAAAGGGG
CAGTCAGTCGGCTTAAAACTGTGACCTGCATGCATTTGCCAACTGCAGACTGCAAACTGCAAACTGCAACTTCGAAGGCAACCCGAAATGCATCTGCAACTGCAACTGCAGC
TGCCGTGCCATTTGTGTTCATCTACATAATCCCAGAGCAGCGTGTCCACGTCACGCTCATCACAACGCGACGGTGCAAAAATGCAGCGAAATTGGATGCGTTGCTCGTTTGT
CAACTGAAATGGTCCAGATGCACAATGCAAGGGAGGGAAGATGCGGCAGAAATGCGCCGGAATCGATGCAGTAATGCGTTTCAAGCTATCATTATGGGATCACATGGGGTCA
CATCAGTGCGAAAATGCTAATGCTATATGGGCTAAAACCAGTTGGGAATTAGAGCGCCGTACAAGCTGTAAGCTAAGCTGATATGCAATACCGCTTATTAAAGAGCATTTAA
ACTATTTGACTTGAAAATATGACTAACAAATTGAGCAGGGAAAATGCTTTTCCTTCGGAAAAAAATGATTGTGTGCCAAAAACTAAAAACGAAAGTTTCCCTACTATTACTTT
TGGCATCGCCCAAAAAAACACTTTCTTTCATTAGCCATTTCACTGATTATCTTTTTTGGTAACCAAGCGGAGTAATCATCATAAGCACACAGCGTTAAACTATTCAGCTATTT
CCAGCTTGTTGAGATAACTAAAAGCTCGGACACGCTTGAGTTCCACTTTTTTTTTTTTTTTTTTATCTCCATCAACATTTAATAATTAAGTGATTTCGCTTATTGTTCTAGT
ATTTATTAGCGCATTACCTAATTGTTTGACACGACGAAACATGAAATGATTACTAATTTCATCATAGAATGCCAGTGACCCATGCTCGTTGGCATTTCGATTATAATAATAT
TCGCAGTTGCATAAATATAGATTGTTTCCGAACGTGACCTGTACTTCCTAGTGCAACTGAATCTTTTGCGTTTCTGGGTCAGCTTTATCTGCCTAGTTCGCTGACATTTCA
AAAGTTTGCAACACTTGGCAGCTTGGCAGCGTAATGAATTCCTTCCTGTTGACTCCGCCACGCAAATTGGCACAAATTTCGCAATAAGAATACTTAATGTCCAACTCCAGAC
GCAATCGCAGCGGCACACGCACACATGCTCTGCGTTAATTTGAGTTTAGCGCCAATTTTTAATTTGAACGTTCTGCAGTTTTTCAACTTATAATTGCACTCTGGCCCATTTA
CATTTTACGCAGGGCATTATTATTATTTCTTTTTTTTTTTTTTAATTTTTGCCTGCTGAGTCGGTCGACATCGGATTCCTCGTTGGAGGCCGAAGGAAGCGAAGTAGGGGGCA
TTGTTGTCATGCATCCACAGTTGCGGTTAATGCAATAGCTCTGAGATGGATAATAATTACCTCGGTCCTTGAATGGTGCGAAAATATCTTGGCAGTTCCTATAAAAGCATAG
CATTGCGTTTGGTACATGAATTTCATTGGCAAATAATAAGTATTTCATCTGGCAGCAGTTCACACCTCAAATCATATCGCGCAAATGCCACCTTTTCGCAAGAATTTGAAAT
CAAAATAAATCTCAGAATTCCTGATTACGCGATGCCTGGAAATCTAAACAGATCATGAGTAAGCCGCAAATGTTTCAGCTAAGAAAAATGAAAAGAACAAAGTTTCCGCTAA
GGGTAGTTGGAAAAGTAGTCAAATTGGTTTACTGATCAATATTACTTTTAGATCTCTCTGTTGTTTTATTTATTTAGCAGTGTAAACTAAGACTGCAAGTCCATGAACTATA
GCCATTTTCAACACTGCTTCTAAGTACTAATGTCTTGAGTGCACTGTACCCACATGCGCACGCATATCATATGAAAATTTGTACTTATTTGTGCCCCTTTTCTTTCTTATTT
TTTTTTTTGCAGACAACTGAGACGACGAGCGACAATGCGAAAGTCAGGCCCGGATACAAAAGTCCAGACCTCGAGCGAGCGAAATAATGGCAATGGGCAGCTCAGAAATAAG
AAAGAAAGAATGGAAAAAACACGAAAAAAAAACGCTGCCAAAGCTGGAGTTTCGAGACGAGCGCGGCACACACTTGCGCAATCGCAGCACAAAACCAAGCTGACCAGCTAA
TTGTGCGAGTGTACGAGCGAATGAGTGTGTGTGTGGTTGTGTGTGTGTGCAAGCGACTTGCGCGTGATTTGAAATCGGCCAACAACCGCTGGGGAACAAAACGTACTCAGT
ATGCCAGTATTCAACTTCAATTTAAGCCAGCGCACAAAAGAAACGTATTATTTTATTCATCATCATCATCATAATCGCATAGCATCGCATCGCATCGCATCGCATCGTATCG
CAGAATAACGGTAAAGCAGTTCGGGTCACATTAACACGTTCTAATTGGACACTCCGGTCAGTGTTGCGAAATGGAATGTGCAGCATAACGAAACAGGAATTCAGCCTAGACA
AAACAAAGCAGCGCTATTGTGGCTCTTTCTTCTTTTTAAGCACTTTTAGCCAACAGAAATAAAAACACGAGGGCTTTATGTGTGTGCATCTGGCTGTATGTATGTATGTGTGGC
AACCACAAGGAATTTTCTGTTGAGGCACTTTTGTTGTTTAGTTTAGATCGCGTTGCTCTGGTTTTTTTTTTTTTTACTCAAATCTCAAGTGTGAGAAAAACAGCAGGCGCACC
ACACACAAACACACACACACATGCATGCACATATCGGATGTTTAACATTTAAATTAGCTGCTTCTGATTGGCATTTTCGGTATGACACTGCTTGCATTCAGTTCTTTATGTG
TGGGTGTGTGTGTGTGTGTGCCTTTGACGTTTTTATTGCCTAATCACTTTTTTTTATTTGCACTTTACTACGAGAATACTGGACGAAAACCGCTGACTGGAACTATTTTTAT
GTTTGCCCCATAATGTTTTTCAACCGAAAAATTGTATATTTTTTTTTTACGTCTAGTGTAAACAATTGAAATTTATTTACTTCAGGGCATAAACAAAATCGTTATACACACA
CATTGATAGGCG
(SEQ ID NO: 22)

Exon: 1001..3033

FIGURE SHEET 13

Exon: 7293..7412
Start ATG: 1001

Transcript No. : CT14539
ATGGACAGGAGTCGGAGTAGCAGAGCTGCCAGCAGCAACCAGTTCATCAGGCCATGTGGCCTACTGACGACCGTCATCCTGCTGCAGACGCTTGTGTCGATGAGTTTGGCCA
TCGAGGAGATGTCGCCACCACCGGCCGCCCCGCCACGTCCATCGCCGCCGCCCACCGTGAAGCTCAACAAGTGCTGCCATTCGGGCGAGTATCTGAACGATGGCACCTGCAT
CGCCGGCAGCGAAGCCCTGTGGCTGCCCATGGTGTACTTGGTGCAGCAGCAGCGCTTCTTCGAGCCCCATGGCGCCAGTCCGCGATTCTTGAAGTTCCTGCCCAATACGAGA
CCCACGTGCCGAAAGGACCAGACGACGGAGATCTTTCGCAGCCGCGGAGCCAATGTGATGCTCTTTCCCAATGGCACGCTCTATGTGCGGGAACGCGCTCTAATGGTCCAGC
CGTCGGATTACTGTGTCGACTGGGAGGTGGCCGTCGTTTGCTTGAACGACAGTCAGCCGATCAATGCTCTGGAGGATCCCGATTATGCAGCCAATCCTCTCGTGCAGCAGGA
GCCGCCCAAGGCTAGCCTGCGACTGAGCAAGTGCTGCGGCAAGTGGGGCAGCTATAACACGCAGCTCCAGAACTGCGATCTGCAGCCCAACCATCAGGCAGCCGTCGATGGA
CTGCTGCGATTGTCGCCGCAGCTACCGGAGGGCAGCTATCAGACTAGCTACGGCCTGCCCGACTGTGGCCAACCCGGTGGCTACTCCATTGCCGGTGATTGGCAGGATGCCA
AGCTGGATCGGAACACGGCCATGCTCCAGTTGCCGCACAAGAACCTCAGCGCCGGACAATACTGCCTGGAGCACACGCAGCGCGAGGGCGAGGTGAAGATCATAGCCTGCCA
GCATTTGTTTAGCTCCGCGGCAGGAGCCGGCATCCACGACGGCTCCATTGGAGGCACCATCGAGCAGGCCAATGGCCAGAATCTGCAGAAGGCCGTCCTGACTGGTGGCATA
CTTGTGTCCATTGTCTTCCTGTCCGCCACTTTGGTGGCCGGCTTCCTGCTGCCCGCTGTGCATCATGCGCTTCATTGGCGCTGCCAAATCTGCTATGTCACCTGCTTGCTAT
TCGGCAAAATACTGCTGGCCATCGAGGAGCTGAGCTCCAGCCTGCAGCCGGGCAGTGCCGCCTGCCATACGCTTGCTATCACCATGCAGTTCTTCTTTCTGGCCGCCTTCTT
TTGGCTGAACACCATGTGCTTCAACATCTGGTGGACGTTCCGGGACTTTCGGCCCAGTTCGTTGGAACGCAATCAGGAGGCACTGCGTCGCTATCTGTACTCCCTGTACGCT
TGGGGCGGTCCTCTGCTTATCACCTTCGTGGCCGCCTGCGTGGACCAGCTGCCGGAGACGACGCTACTGCGTCCGGGATTCGGACAGCTCTACTGCTGGTTCGACAATCGCA
ATCTCTCGATCTTCGCCTACTTCTACGGACCCATTGGCCTGCTGCTGTGCGCCAATATAGCGCTCTTCGTGTCCACCACCCATCAGTTGACGTGCGGCCTGTGGAAACGGGA
CGATGTCAAGTCCTCATCGGAGAAGTCCGCCCTGGGACGCGTCTGCCTGAAGCTAGTCGTCGTTATGGGTGTCACCTGGATAGCGGACATTCTGTCCTGGCTAGTGGGTGGA
CCGCATGGCGTATGGTTCTTCACCGACCTGATCAACGCCCTGCAAGGCGTCTTCATCTTCATCGTGGTGGGCTGCCAGCCGCAGGTGTGGACCGCCTGCCGCAGGATCTTCT
GTCCGCGATTGCGTCATGACATCACCAATACGACCAACGGTGTCCAGCATTCGAGCAGCTCGCAGGGTTTGCCCTCGATGGCCGGCGGCACGGAGATTACACAGAACACCAC
CACCACCACCACGACAACCAACACGACCGCCACTCACATGCCCAGCAATCCGGCGGAGGATGAGGTACCGGAGAAGGCTCCCATTGCCCCCGTCGCACCCATCGTCAAAATG
GAGACCATTTGCTAAACACAACTGAGACGACGAGCGACAATGCGAAAGTCAGGCCCGGATACAAAAGTCCAGACCTCGAGCGAGCGAAATAATGGCAATGGGCAGCTCAGAA
ATAAGAAAGAAAGAATGGAAAAAAC
(SEQ ID NO: 23)

Start ATG: 1

MDRSRSSRAASSNQFIRPCGLLTTVILLQTLVSMSLAIEEMSPPPAAPPRPSPPPTVKLNKCCHSGEYLNDGTCIAGSEALWLPMVYLVQQQRFFEPHGASPRFLKFLPNTR
PTCRKDQTTEIFRSRGANVMLFPNGTLYVRERALMVQPSDYCVDWEVAVVCLNDSQPINALEDPDYAANPLVQQEPPKASLRLSKCCGKWGSYNTQLQNCDLQPNHQAAVDG
LLRLSPQLPEGSYQTSYGLPDCGQPGGYSIAGDWQDAKLDRNTAMLQLPHKNLSAGQYCLEHTQREGEVKIIACQHLFSSAAGAGIHDGSIGGTIEQANGQNLQKAVLTGGI
LVSIVFLSATLVAGFLLPAVHHALHWRCQICYVTCLLFGKILLAIEELSSSLQPGSAACHTLAITMQFFFLAAFFWLNTMCFNIWWTFRDFRPSSLERNQEALRRYLYSLYA
WGGPLLITFVAACVDQLPETTLLRPGFGQLYCWFDNRNLSIFAYFYGPIGLLLCANIALFVSTTHQLTCGLWKRDDVKSSSEKSALGRVCLKLVVVMGVTWIADILSWLVGG
PHGVWFFTDLINALQGVFIFIVVGCQPQVWTACRRIFCPRLRHDITNTTNGVQHSSSSQGLPSMAGGTEITQNTTTTTTTNTTATHMPSNPAEDEVPEKAPIAPVAPIVKM
ETIC*
(SEQ ID NO: 24)

Name: mth-like 1
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384553
GACCAGACAAGTGCAAGCAATGGAGCTCAGCGTCAGTATCCACTGTCTCACACACATCCACAAAAGCATCTTCATCAGATTCAGTTCCTCAGATCTCGCATCAGCTGCTGCG
ATTCCCATGGACGTGACACGGTGGAGCGGTGGAGAAAACATCATCATAATAATAATAAAGCCGAATAACTTTCTCCACACGTAGTGTGTGTCAATAAATCGGTTGTTTCTCG
CAGAGTTTGAAAGGAAACTTTTAGAGTTCTTAAACAAAGAATGTGGTTTTGGAACCACTGTGCCGCATCCTGTCTGCCATGAATGATGCGTCCTTTCCTGCCAGCACATAAA
GCCCCGGCCAGTTGCATCTATTCCACACACTGATAGACAATGTTTGTGCACTGATTTATACTCAGCCAGCTATTAAATAAGAATTTTCGTTATTACGCTGCCGGGGATTTTA
TTATGTGACTTTCAAAAGCCCCCGCCAACTTGTTGTGGCTTTGGCTTAAAGCCGCGCCAAGCTGAAAATTGTTATTTTAACCTTTTAAAGTCAGCTTTATCAGCACCAAATG
CAGGGCCATTATGCGTGGCCCACGCGGCGTATGAGCAACGTTAAATCTGCGAATAATTCAAAGACGTGCGTTTCTTTTTACAGCCGCCTTGTCGGTAATCCCATTAAGGAGT
TATCTGGCGAAACATTTTTGCACAACACTCGACTGGAGGCACTGTGAGTATACTTTGTTTGTTTTAATTTAAAAAATCACAAAATATATCATGAAGTTGTTAAAACGGGATT
GAATTTGATTTGAATATATGATTTCATTTAATTTCTCACCAATATTTAAGTTCGTTGGCCCTAATGCCGATTCACATCAGCAGCAGTTTAATGGAGCCACTGAATATATCGT
TTTTGAATCTCACGGGAATTCGGTATGATCACATCGACTTCGAGGCCATCAATTCAATGCGCAACCTGACGTACATCATCTACGATCGATTCTTCTACTGCTCAATGACGCC
GCGGGTGAGAATGTGCAAGCCGTCCACCGATGGAGTCAGCTCCTTTCAAGATCTTCTGAGCAAACCAGTGTTACGTTACTCCGCCTGGGTGATGGCCACGTTGACAATAGCT
GGAAATGTGCTAGTCCTATGGGGACGCTTTATCTACCGCGATGAGAATGTGGCTGTGACAATGGTGATTCGGAATCTGGCTCTGGCTGATATGCTGATGGGCTTCTATTTGG
TCACCATTGGAGTGCAGGACTATCGCTACCGCAACGAGTATTACAAGTAAGTCAAATGTGATGGATTGGCATTTGTTACGCCCCAATCTTTTAGGGTTGTGCTGGACTGGAT
AACCTCGTGGCAGTGCACACTGATTGGAACGCTGGCCGTGAGCTCCTCGGAGGTTTCGATGCTTATCCTGGCTTTTATGTCGCTGGAGCGGTTTCTCTTGATCGCCGATCCC
TTTCGCGGCCATCGCAGCATTGGAAATCGGGTGATGTGGCTGGCTTTGATTTGCATTTGGATTACGGGCGTGGGCTTGGCTGTGGCACCGGTGCTCCTGTGGCGGACCAGTA
CCCTACCCTACTACGGATCCTATTCCGGAACCTGTGAGCACAATGTTTATGATAAAATGTAGTTTACCAAGTGACGAAATTAAACTTAACAGGCTTTCCATTGCACATCCAC
GAGGCCTTCCCGATGGGCTGGCTTTACTCGGCTTTCGTGTTCCTGGGCGTCAATCTGCTGCTGCTGGTGATGATTGCGATGCTCTACACGGCACTCCTTATTTCGATATGGC
GAACGAGGAGCGCCACTCCGCTCACTCTGTTGGATTGCGAATTCGCTGTGCGATTCTTTTTCATCGTGCTGACCGACTTCCTGTGCTGGGTGCCCATCATCGTTATGAAGAT
CTGGGTTTTTTTCAACTACAATATCTCCGATGACATCTACGCCTGGTTGGTGGTCTTTGTGCTGCCACTTAACTCGGCGGTTAATCCTCTCTTGTACACCTTCACCACGCCC
AAGTACCGAAATCAAATCTTCCTTCGCGGCTGGAAGAAGATAACTTCTAGAAAGCGAGCGGAGGCTGGCAATGGAAATGTGGCCACCACAACGACGGGCACGGCAACTGGAT
CCTCTCAGCATCCGGATGACTTCACGATCTTTGCCAAAGCCGCGATGAGATGCCATTAGGTCCTAACGCTGTCGAACTGAAGGATTCTTGTGGCGCAGATCGAACTTGTATA
CCACCAGATTTAGCAGCATTTATTTTGTCTATGATATCGAAAATCTATATATGTATGTGAATATCTATCCCTGTCCGTGTAAGATAACCAAATGAAAGTGTGAAATAAATTG
AATGTGATATTTGTAAGGCAATTTTACGGAATTTAAAATGGGATGCGAGTTAGTGTAAGATTAAGCTATTATCTAAGTTTTGGTAATAGTTTTAAATACTTTTTATTTAGAA
ATAGCAGCGCTATCATTTTTGGCATGGTGTGATTCTTTCTGGTCCGAGTAAATCTTTGCCATTACCAACTTGGAGCGCACGATGTGAAACTTTCGCATTTCCTCCAGGTCCT
CCATCTGCACCACCATGTTATGCTTACGCTCCAGGTGCATGCGTAGTCTAAAAAATTATTAAGGAGTGTTACTTTCAAATAATGTACACTTTAAACAAAATAACTAACTTGT
TACGAGCTGGATAGCTCTGATGGCAGTGGGCGCACTCAAATGGCATGGCCGTATGCGTGCACTCGTGATTCCTGAGAGATGAAGATGAGGTAAAGGACTTCTGGCAGCCGGG
AAACTCGCAAATGAAGCGAGATCCAGAGTGCTTAACCTGCTGGTGGACAGTCAAATGGGATCGCTGTGCGAAACGACGATCACAGAGATCACAGGAATAGGGAAAGTCTCTG
TTATGATAGCGTTTATGAATCTCCAGGGCTTTGGCAATGGCAAAAACCTTGCCGCATGTGTCGCAGTGGACTCCAGACTGCCTGTGCTTCGAATCGACGTGAACTTTCAGCG
CCTGCTTTGTGAGGTAAACCCGGTCGCACTGGCTGCAACGAAAGTTGGGACGATTGGCCAAATGAATCATCTGGTGCTTGGCCATGCGGGAATGGTTCCTGAACTCCTTGTG
GCAGAACTTGCAGTTGGATTTTCCCGTATGCCGATTGCGTGGCACAATGTC
(SEQ ID NO: 25)

Exon: 1001..1278
Exon: 1327..1601
Exon: 1661..2187

Start ATG: 1001

Transcript No. : CT16185
ATGACGCCGCGGGTGAGAATGTGCAAGCCGTCCACCGATGGAGTCAGCTCCTTTCAAGATCTTCTGAGCAAACCAGTGTTACGTTACTCCGCCTGGGTGATGGCCACGTTGA
CAATAGCTGGAAATGTGCTAGTCCTATGGGGACGCTTTATCTACCGCGATGAGAATGTGGCTGTGACAATGGTGATTCGGAATCTGGCTCTGGCTGATATGCTGATGGGCTT
CTATTTGGTCACCATTGGAGTGCAGGACTATCGCTACCGCAACGAGTATTACAAGGTTGTGCTGGACTGGATAACCTCGTGGCAGTGCACACTGATTGGAACGCTGGCCGTG
AGCTCCTCGGAGGTTTCGATGCTTATCCTGGCTTTTATGTCGCTGGAGCGGTTTCTCTTGATCGCCGATCCCTTTCGCGGCCATCGCAGCATTGGAAATCGGGTGATGTGGC
TGGCTTTGATTTGCATTTGGATTACGGGCGTGGGCTTGGCTGTGGCACCGGTGCTCCTGTGGCGGACCAGTACCCTACCCTACTACGGATCCTATTCCGGAACCTGCTTTCC
ATTGCACATCCACGAGGCCTTCCCGATGGGCTGGCTTTACTCGGCTTTCGTGTTCCTGGGCGTCAATCTGCTGCTGCTGGTGATGATTGCGATGCTCTACACGGCACTCCTT
ATTTCGATATGGCGAACGAGGAGCGCCACTCCGCTCACTCTGTTGGATTGCGAATTCGCTGTGCGATTCTTTTTCATCGTGCTGACCGACTTCCTGTGCTGGGTGCCCATCA
TCGTTATGAAGATCTGGGTTTTTTTCAACTACAATATCTCCGATGACATCTACGCCTGGTTGGTGGTCTTTGTGCTGCCACTTAACTCGGCGGTTAATCCTCTCTTGTACAC
CTTCACCACGCCCAAGTACCGAAATCAAATCTTCCTTCGCGGCTGGAAGAAGATAACTTCTAGAAAGCGAGCGGAGGCTGGCAATGGAAATGTGGCCACCACAACGACGGGC
ACGGCAACTGGATCCTCTCAGCATCCGGATGACTTCACGATCTTTGCCAAAGCCGCGATGAGATGCCATTAG
(SEQ ID NO: 26)

Start ATG: 1

MTPRVRMCKPSTDGVSSFQDLLSKPVLRYSAWVMATLTIAGNVLVLWGRFIYRDENVAVTMVIRNLALADMLMGFYLVTIGVQDYRYRNEYYKVVLDWITSWQCTLIGTLAV
SSSEVSMLILAFMSLERFLLIADPFRGHRSIGNRVMWLALICIWITGVGLAVAPVLLWRTSTLPYYGSYSGTCFPLHIHEAFPMGWLYSAFVFLGVNLLLLVMIAMLYTALL
ISIWRTRSATPLTLLDCEFAVRFFFIVLTDFLCWVPIIVMKIWVFFNYNISDDIYAWLVVFVLPLNSAVNPLLYTFTTPKYRNQIFLRGWKKITSRKRAEAGNGNVATTTTG
TATGSSQHPDDFTIFAKAAMRCH*
(SEQ ID NO: 27)

Name: TSH, FH, LH receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013383791
CACTCAGATTTGCCGACAGCAGAAATATTTTAGGGTTGTATCTTAGAGGTGGAGAAACATCGATGCCAACATCGATGTCATCGATGTTTTCGAAAAATCAAAAACGTCGATG
TTTTTCGAGGAATCGTCGATGTTTTCAGCGCATAGAAACAGCAAAGTGTGTAAATTGTAAGTCTAAATTAATTTCCATATTTTCTAATACAATTATTAATAGCATATTGCAT
ATTTATTGTGACTCGCCGGTGGTGTAGGCTGGTGTCTTGGAAGACATGGATAAATATCTGAAATGTAAGTTTTTTTTTTAAAGATCAATTTATTTGTTTTTTGCCAAAGACA
TCGAAAACATGGATGTTCGCCAACATCGATGTTTCTCCACTTCTAATGCCAACGAAAACGTGACTGGGAACGTATCGATGGGAAAAAACGTTACAAAAATACCAACTATCGCA
TAGCGGAATCACCTTGGCGGGCTATTCGAATGTATCTCACAAATAGTTTTTAAGGAACCTACAGTCAGTACCAAATACTTAAAATTCCATTTATATTTTATTTTATGCTACA
GAAAACTGATAAGCTGCATTACTTTTGAAAAGTTAAAAATTGTGTTTTTTAAATGATTTATTTATATAAATAAGCTTATGTAATACTATTGTTAATATGTAAAAAAAAAAGA
ATAAAAGAAAAAGAATAGTTTACCTAAGCAGTTGCTTGAATAATACATACATAATTTAAGTTTTATACATAATTTATCAACCCTCGCTCAGTAGGAAAAAAAGATGGCTCAG
AAATGAGCAACGAATTGGAATGTCCGATAAGCCTTTTTTATCAAAATAATAATTTTTTACATTCCCTTTTCATAACGCTTTCCCAGTTAGTTTGTGCCAATCGTAAAACAGCA
TCCGCAAAGTGAAGTCGCTGCTATTTTGTAGTACCACTACATTCTTATTTATTTAAGTAATATTTAAAGAATGTCCAGCTAACTATATGATTGCTAGTTCAAAAATGTTACT
TTCTGCATCGATCCTAATATATTTCTTGCTGAATCTGCAATCTTCGAGTGCTGAGATTGCCGACTGCAGTTTCTATGATACCGTTGATATCTCGGAAGGCCAGAGGCTCTCG
AATGGATCATACCTCTACGAGGGATTACTTATTCCCGCCCATTTGACGGCCAAGTATGAATTCAAGCTTCTGGCGAATGGAGATAAGGAGCAGGTGCCAAGTCACGTGAGAG
GATGTGTGTGCAAGCTAAGGACTTGCGTCAGGTTTTGCTGCCCACATGACCACATAATGGATATGGGCGAATGCTACGCCAACATGACGACGGAGGAGAACGAATTACTGGA
CCCCATGCTAAATGTGACCCTGGACGACGGATCAGTGGTGCAAAGGCACTATAAGAAGGAGCTCATGGTGCAGTGGGATCTGCCAAAGCCCTGCGATGATATGTTTTACCTA
GACAACCGGGATATAATGGATGAGTACACCTTGTTCGAGGTATGGCTGCTCGAAAAATAAGCATTTTTGGGCAGGATTTATTTATGCTTTGGCTCTCTGCAGAATGGAAGAT
TGCTGCGCCACTACGATCAAGTGTATTTGGACAAGTCGGAGTACTGTCTGCAGCACCGAACATTCGGCGAAGGTAACAACAACTCCATTCGCATTATACCCCACAACTGCTT
GATACTGCCATCCAGAACGGGACAGACCGTTGGTAAGATGATGCCCATTGCTTGCATGGTAGCTCCCAGTTCTATATCTTTCCTATTTCCATAGTGATGATCACATCGTTGA
TATGCTTGGTTCTTACCATCGCCGTGTACCTGTGCGTTAAGAAGCTAATGAATTTGGAGGGCAAGTGTTTCATCTGTTACATGATGTGCCTATTTTTCGGATACCTGTTCCT
ACTCCTTGATCTGTGGGAGTTATCGCTCGATTTCTGTAAAGCAGCAGGTTAGCTTAGAGGGGATGCATTTCAACACAACTTGACAAACCATTGACCATTTACGACAGGTTTC
TTGGGATACTTTTTCGTCATGGCTGCGTTCTTTTGGCTATCTATCATCAGTCGGCACTATTGGAAATGTCTTACCAATCCCTGTGCCTCGATGAACATTCGGTCAGAACGTG
CGTTTCTGCTATACAGCTGCTTCGCCTGGGCAATGCCTTTGGCCTTGACCGGAGTCACCTATCTGGCTGATAACGTTGTGAACAACGAGGAGTGGCAGCCTCGCGTGGGCGA
CGAAGGGCATTGCTGGATCTATAGTGGGTTCTGTTTTGTATCCGGCTTGTGGTACCAGCTCACCAATGTTTTTCTTTCCACCCTCAGCTAAAAGCTGGTCTGCAATGGTGTA
CTTCTATGGACCGATGGTGCTCCTGATTCTCTTCAACATAACTATGTTTGTTCTGACGGCTAAACACATAATAGATTCCAAGAGGACTCTTCGAAAGATTGCTCGCAATGAA
GGTAGAATACAGAAGCTGAATTCGGACAAACAGAAGTATATCAATTATAATAGTTGAAAATGTCAGTTTAAGATAAATATTATCAATTCGTTACAGTTACACACAATTTCTG
CTGCTCTTCACTGTGATGGGGATGTCGTGGAGCTTTGAGATATTCTCCTACTTGGTGCAACGCGAGAAACTGTGGGTCAATATCTTTCTGGTAGCTGACTACTTCAACTGGT
CCCAAGGCGTCATTATATTTGTGCTATTCATTCTGCGGCGCAAAACTCTTGTACTATTCAAGAAACAGTAAGTCTTATAGTTCATAGTTCGGGTAACCAGTTCCCTAAAATG
TCTCTTTTCAGAATATTTCCCAAACAGAGAGCCTTTAGTAGATCTGCAACTCAAAGTACGATTGAATCCATTTCTCAAACAAAAAGGCATTTTAACATGACATAGTATTTTC
TTCTCTTCGAATTAATCAATTTAATAGTAAAGTATAACTTTTTTTTCTCACAAAATGTATGAATATATTTCACTTCATTCTACATTCTATATAGAATTCATTTAATCACAAAA
TGTATGAATATTTTTCTTTATATGACGTGCATCGACTATAGAAAATCGACTATAGTGTGCCACAATATATTTTACTCATCATTTCCTAATCAAACTATTGCAAACTTAACTT
TGAGCCGAGGAAAAACCCTGTATGTCTATCAAAACTATGAGTCAAAAGTATGACGCGATGCTGAAAGTAGTTGCCTTTAAATATTGAAGTGATAAATGTAGAAAGAAGACAT
TTGTATGGAAATAAACTAAGTTTGTAAATCTACATTCGCTTTATTTCATGTACAAATTTTAATTTAAAAACTTTGAAGGTAAATAGTGCTTTATCAGATGCGGATTACACGG
TAGCAGTTATCATCGATGTCGAATCTTCTTGTTTATGTGACCGCTTTTCTTCGCTGTCGCACTGGAGTTCGTGGATAGTTTGCTCTTTGTATTCGAAGTGGGTGTGCCGAGG
AGATCCTTGTTCCCTCCACTGTTCGTAGATCCAGACACCATTCCGCTCACTACGCTGCCCTCATTGCGATGCGAGGGTGAATCTGTAAATAAAAAGGTGTAATTAAACATAT
ATACTTTTCTGGAGAAAAGCGCTAAAGTTGAATCTCAATGTGCTGTGCTGCCTTACCCTTTTCGTCTTTCGGTGTGCTCTTGATTGTCGTGGTTCCGTCGTGGGCCGTTGTG
GTGTTCTGATTGATTGAGATCTGATTGTCGCTGGAGTCTTCATTTGTGATGGAATCCGATTCG
(SEQ ID NO: 28)

Exon: 1001..1495
Exon: 1559..1712
Exon: 1775..1951
Exon: 2012..2263
Exon: 2328..2499
Exon: 2561..2759
Start ATG: 1001

Transcript No. : CT16507
ATGTTACTTTCTGCATCGATCCTAATATATTTCTTGCTGAATCTGCAATCTTCGAGTGCTGAGATTGCCGACTGCAGTTTCTATGATACCGTTGATATCTCGGAAGGCCAGA
GGCTCTCGAATGGATCATACCTCTACGAGGGATTACTTATTCCCGCCCATTTGACGGCCAAGTATGAATTCAAGCTTCTGGCGAATGGAGATAAGGAGCAGGTGCCAAGTCA
CGTGAGAGGATGTGTGTGCAAGCTAAGGACTTGCGTCAGGTTTTGCTGCCCACATGACCACATAATGGATATGGGCGAATGCTACGCCAACATGACGACGGAGGAGAACGAA

TTACTGGACCCCATGCTAAATGTGACCCTGGACGACGGATCAGTGGTGCAAAGGCACTATAAGAAGGAGCTCATGGTGCAGTGGGATCTGCCAAAGCCCTGCGATGATATGT
TTTACCTAGACAACCGGGATATAATGGATGAGTACACCTTGTTCGAGAATGGAAGATTGCTGCGCCACTACGATCAAGTGTATTTGGACAAGTCGGAGTACTGTCTGCAGCA
CCGAACATTCGGCGAAGGTAACAACAACTCCATTCGCATTATACCCCACAACTGCTTGATACTGCCATCCAGAACGGGACAGACCGTTGTGATGATCACATCGTTGATATGC
TTGGTTCTTACCATCGCCGTGTACCTGTGCGTTAAGAAGCTAATGAATTTGGAGGGCAAGTGTTTCATCTGTTACATGATGTGCCTATTTTTCGGATACCTGTTCCTACTCC
TTGATCTGTGGGAGTTATCGCTCGATTTCTGTAAAGCAGCAGGTTTCTTGGGATACTTTTTCGTCATGGCTGCGTTCTTTTGGCTATCTATCATCAGTCGGCACTATTGGAA
ATGTCTTACCAATCCCTGTGCCTCGATGAACATTCGGTCAGAACGTGCGTTTCTGCTATACAGCTGCTTCGCCTGGGCAATGCCTTTGGCCTTGACCGGAGTCACCTATCTG
GCTGATAACGTTGTGAACAACGAGGAGTGGCAGCCTCGCGTGGGCGACGAAGGGCATTGCTGGATCTATACTAAAAGCTGGTCTGCAATGGTGTACTTCTATGGACCGATGG
TGCTCCTGATTCTCTTCAACATAACTATGTTTGTTCTGACGGCTAAACACATAATAGATTCCAAGAGGACTCTTCGAAAGATTGCTCGCAATGAAGGTAGAATACAGAAGCT
GAATTCGGACAAACAGAATTACACACAATTTCTGCTGCTCTTCACTGTGATGGGGATGTCGTGGAGCTTTGAGATATTCTCCTACTTGGTGCAACGCGAGAAACTGTGGGTC
AATATCTTTCTGGTAGCTGACTACTTCAACTGGTCCCAAGGCGTCATTATATTTGTGCTATTCATTCTGCGGCGCAAAACTCTTGTACTATTCAAGAAACAGTAA
(SEQ ID NO: 29)

Start ATG: 1

MLLSASILIYFLLNLQSSSAEIADCSFYDTVDISEGQRLSNGSYLYEGLLIPAHLTAKYEFKLLANGDKEQVPSHVRGCVCKLRTCVRFCCPHDHIMDMGECYANMTTEENE
LLDPMLNVTLDDGSVVQRHYKKELMVQWDLPKPCDDMFYLDNRDIMDEYTLFENGRLLRHYDQVYLDKSEYCLQHRTFGEGNNNSIRIIPHNCLILPSRTGQTVVMITSLIC
LVLTIAVYLCVKKLMNLEGKCFICYMMCLFFGYLFLLLDLWELSLDFCKAAGFLGYFFVMAAFFWLSIISRHYWKCLTNPCASMNIRSERAFLLYSCFAWAMPLALTGVTYL
ADNVVNNEEWQPRVGDEGHCWIYTKSWSAMVYFYGPMVLLILFNITMFVLTAKHIIDSKRTLRKIARNEGRIQKLNSDKQNYTQFLLLFTVMGMSWSFEIFSYLVQREKLWV
NIFLVADYFNWSQGVIIFVLFILRRKTLVLFKKQ*
(SEQ ID NO: 30)

Name: mth-like 2
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384913
ACGTTGTCAGTGTGTGTGACGGTTGTTAATTTCATTTGACTTGCGAGGGTTTTAATATTGCAATGGAAATTCGGTACAATTTGCCGTCCATTAAATAAGCACGCAATTAATT
GTGGCAGCCATGACAGGCGTGGCTTGCCACCAATTAAGTGGAATACTATATGGCTATATCGCTATATCGCTATATGTCTATATGGCATATAGATGCCCACAGCCCCTTGGGG
TCGATCGCCTACGCCGTCTATAATCCCCAGCTAAACTTCGAGACATCGCATCCGGCACCGCAATCCGAAATCGGAATGCTGATGAGGCAACAGAGTGGACGCGGCCGGAAAC
AAGGCACTCCTCCCATCAATCTGTCAGACGGGCCAGTGAGTCCTGGATGTCCTGGTGAACGCAGCTCAGCTCCCAGCGGGTGAGGCCAGTTGCCCAGACAGTTAGTCCGCCA
GGCAGTCAGTCGGTCTGTCAGTGAGTCAGTTAGTCAGTGAGCGAGTGAGTCAGTCAGTCAATCGAACAATCAATCATACGGAATGACAGCTACGTCGGCTTAAGTCAAAGGC
AGTCGGCCAAGAAGGGATCTTGTCTAACTTACAGTCGCAAAATTCAATAGCCTGTGGCAATCGCACTTGCAACTTTAATAATTCCAAGAACGAAGCAGTCATAATGTTTACT
GAAATAATGTTATGCCACAATTGGCACATACAAGCTCATAATAGTGTTTTATTCTGTTTTAATTGCACTTTAAAATTGATGATTGTATTAGCCATCATAGGCATGGCTTGA
GCTGTTCATTTAGCCATTTGATTTTGTCTACGTTGGTTTTTTATCAAGTTTCTTGCCGTGTAACCTTTTTTTTTTTTTTGGGCTCTCCCAACAATTTGAGCTCCTTTTTCGT
TTTCGCATTCCCTTAAACATTTGCTGGGGAAACACATTGATTTATATGCCGGCATCGCCAAAAATTGAATGTTATTGCTTGAATGGCCCCAATTTACTTAAAAGCTACGCTC
CTCCGCCCCACTGCCCCTCTGCCGCTCCACTGACCCCTGACCCCATTATACATGTTGTGGGCTCCACTGGCTCCTGATTGGTATTTATGAGCGTGTTTATTTATTATATGAC
CCAACACTTGATAAAATTAAACGGAAATGGTTGGCAATTTTGCATTTGATTCTGACACCCAAACACACAGGCAGGCAAATTCGCAAATGAACTCATCTGCCCAATATTCCGA
CCATAAAGCATCCCGACCAACGGCACTCACATATAAGTTAGTTAAGTATGTATATTAGTGGCTTGGCAACTTACCAGAAAGTAGACAGCAAATGGCATCACGACCACGGCGA
CCAGGAGATCGGCGATGGCCAACGAGACTATAAAATAATTCGTGACCGTCTGCAGGGAGCGCTCGCGGCACACGGACAGGATGACCAGGATGTTGCCGAAGAGGGTCAGGAT
GGGGAACAGAATCAGGATGAGCGCCCAGTAGTTGTGGTCCACTGAAACGAAGCGAGTTGTAGTTACATTATATCCGGGATGCATCCATCCGCAGGCTCTCCTGCTGCTGCCA
CCCAAATTCCCTGGCCAAAGTTCACCTTGCAGGAAGATTTCAGGGAAAGCTCAACGGGCACGTAAGTCAATTACACAGGCAGTCCGTTTAATCAAGGACTCTGGCTGACAAC
TAATTTTAATGGCAACACAAGTAAGGACTAGCACACAAAGACGTGAATCTTAAAGGCTGTTTTAATATAATCTATCTAATATAGAGTGATGCATGTTACCAACAGCTGCGTT
GTTAAAAATTAAAGTTATTTCACATATTAAAAATGGCAAAATATGGCTATGTAATATGAAAGTTATTAAACTTGAAATACCTTTATCTTCATTATGAAACAAATTAAACAAA
CAAATATGTAGGAACTTTTCTTGAGATTGCACGAATAAATGAGTTATGTAACGCGTTTAAATTAAATTCTTGCTATCAAAACTGATTTCTAAATGAACATCTTTATAAATAA
ATTAATTATCGGCGGTTTCATTCAAATACTCCCATCTTAGAAATCTAAGGAATTCTGTTCTCAAAACCTAGTCACCTCACTATCCGATGAAAATATATTATACTAAAGAAGT
TGCTGTTGCAAGGTTTAAATGATGCTCCAACTTGATGATAATGTTTAAAGGTATTTAATGAGATTCAAGGCTATCTAATCTAATTTAATAATTTTAAAGCCTTTGAATGATA
TTCAATAATATCTAATGAGCTTTAATGATTTTTAAAGATTTATAATGCGGTTTTAAGATGTTAAGTTGTAGTAGTTTGATGATTTCGAATGGTAATCCTGCCTCGTACTCAC
CCACTCTCAGCTCGCCGCAGGAGTCGGTTAGGTTGAGCGTCTCGTTGGTGCAGTTCAGCCGGTAGCCGTCGAGGTACAGTCCCGTCAGGTTCTGCTCGAGGAGCAGTCCGCT
GCCCCTGCTCCCCGTGATGGCACTCCCGGCGCAATGGTGCTTGTTCCCGAAACCGTGCTCCCATTGTAGTTGGGAAAATAATCGTTAAAGTTTTCGAGCAGCAGCATGAAG
GTGCTGCCCCCTGATCCTCCAGCGCCGCCTACACCTCCGCCCGCCGTACCGGCCAGTCCGCCAGCCACGCCCACATCCCCAGAGCCGCCTCCATCGTAACTGGTCGTCGAGG
GCGCCGTTCGATTGCCCAGCGTCTCGCGCGCCACCGTTGCTGTCCGCTGGAGCAGTTCCGCATCACCGAAGCCGCCCAGCATGCCATAGTCCGTTTCCGCATCAGCCACATT
GTCCAGGGCTGCCTCCTCCCCGCTGTCGATCAGCGCGATGCCAGCATCGCCGCCGTCGCCCGCTCCTGCTGCCCCTCCGCCCGCCAGCTGTTGTTCCACAAAGCTGCGCAGC
GAGTAACGCCCCCGCTCACCATCCTCCGCACTGGAGTTGCTCTGGGCTGGCTGCGTGCTGTTCAGTGCCTCCAGCCAGGCCAGGGTTCCGTTGGCGCCCGTCACCGCCAAGT
TTGTGCTGTTCTCGGGCCCGCCCACTGCCGCCTCACTGCCCACCGAACCCAAGAGCCATTCGCTGATGTCGTCGATCAGCGCGTAGCCGGCGTCCTCCGCCGCCGTTTCCAC
ATCCTCGTAGAGCGGCAAGTAGCCGGCTGGCATTGTGGGGCCCTCAACGCCCGCCTGCGCATCGCTGCCGTTCGCCGGTTGAGAGCTGTCCGCTTCCACCGCCAGGTGGTGA
TTTATCCACGTGGCTCCCGCGGATGGGGCTGCCGTGGTCGAGGTGGAAGCAGCGGCTGTGCTCAAGGTCGCCGATGTGGTTGCTGCCGTTGCTGCAGTTGCCGCAGTTGCTC
CCGTTGCCGCCGCAGTTGTTGCCGCCGTGCTGCTCAGTGGCTGGGCTGCCATGGTGCCGCCGGTTCCGCTGCTGCTGATGCCCCGCCGCCAATCGAATGGGCTCAACATCTA
GTGGGGCGGGGGCGTGGCAGCGGAGGCGAGCTGCAACTCCTGCAACGCCAGCGGCATGTTTCCGTTTCCTGCAGCGAAAGAGACGAAGTTCGGAGGTGAGAAAACATAAATC
ACAAAAATTTCCATTATAATCAAAGTTCTTGAAGATAGTTCTTAACGATACGCACAACTAAGATTACAATTTTAAGCCTAATTGGTCGCGTTAATGTTGCTTAATTTGATAA
ATTCTTGTAATCCGCAAATGAGAGTCATGGTTCTTTTCGACTTAATATTTTTAGGATATCTAAAAAACAAAGTCGTTCTCAATTTTTCCTACTTACACATTACACATTGCAT
GCCCCTTCGTTTGCATTTTCAACACAAGGCGTACTTCATGCTTGACCACGCCCCCCCAACCAACACAACAACATACCATCCAGCACCACGCCCCCATGCCATTTTGGACGTA
CGGACCCTTTGCCCAACTGCCATGGCCGCATGTTTAATTAACTCTTTAATCAATTTATTGTCAATTATTCTACAAAAAGCGACAGCGAAAATTGTGTAATTTAATTAAATTA
GCAAAAGGTTTTTGTGTCTCTCTCGCCCCACCAACCATTCCGCCCACTCCCACTTCCATTCTCGCAACCAAAATGCCACGCCCCCATGCTCGACGCTGTCTTGCTTTTACAC
ATTTTGCAATTTTCGCTTTTGTATTTGCCAATTGGCGCTGATTAAAGCCGCCCGAGTCGATGGGCCATAATTAAAAGATGGCCTAATGAGAGGCGAGGCGAGTACGTCTCGA
GTTGGCTGGGGGTCAAAGGTAATCGGATCGGAGTGGTCTTAAGTCCCCTGCCTTCTGCCTCTGTCTTCCGGAGATCGGGTTATTCGAACTCTGAGCTTGTGGATTCTGGACT
CTGAATTCTGGATTCCGGATCGGAGGCAGAGATCCGAGAGCTGCACACCCAGACGGCTGGCACTAATCCAATTTGGGCACACTTCCAGGCGACGAGAGGCA
(SEQ ID NO: 31)

Exon: 3469..2353
Exon: 1497..1307
Exon: 1081..1001
Start ATG: 3469 (Reverse strand: CAT)

Transcript No. : CT17758
ATGTTGAGCCCATTCGATTGGCGGCGGGGCATCAGCAGCAGCGGAACCGGCGGCACCATGGCAGCCCAGCCACTGAGCAGCACGGCGGCAACAACTGCGGCGGCAACGGGAG
CAACTGCGGCAACTGCAGCAACGGCAGCAACCACATCGGCGACCTTGAGCACAGCCGCTGCTTCCACCTCGACCACGGCAGCCCCATCCGCGGGAGCCACGTGGATAAATCA

CCACCTGGCGGTGGAAGCGGACAGCTCTCAACCGGCGAACGGCAGCGATGCGCAGGCGGGCGTTGAGGGCCCCACAATGCCAGCCGGCTACTTGCCGCTCTACGAGGATGTG
GAAACGGCGGCGGAGGACGCCGGCTACGCGCTGATCGACGACATCAGCGAATGGCTCTTGGGTTCGGTGGGCAGTGAGGCGGCAGTGGGCGGGCCCGAGAACAGCACAAACT
TGGCGGTGACGGGCGCCAACGGAACCCTGGCCTGGCTGGAGGCACTGAACAGCACGCAGCCAGCCCAGAGCAACTCCAGTGCGGAGGATGGTGAGCGGGGGCGTTACTCGCT
GCGCAGCTTTGTGGAACAACAGCTGGCGGGCGGAGGGGCAGCAGGAGCGGGCGACGGCGGCGATGCTGGCATCGCGCTGATCGACAGCGGGGAGGAGGCAGCCCTGGACAAT
GTGGCTGATGCGGAAACGGACTATGGCATGCTGGGCGGCTTCGGTGATGCGGAACTGCTCCAGCGGACAGCAACGGTGGCGCGCGAGACGCTGGGCAATCGAACGGCGCCCT
CGACGACCAGTTACGATGGAGGCGGCTCTGGGGATGTGGGCGTGGCTGGCGGACTGGCCGGTACGGCGGGCGGAGGTGTAGGCGGCGCTGGAGGATCAGGGGGCAGCACCTT
CATGCTGCTGCTCGAAAACTTTAACGATTATTTTCCCAACTACAATGGGAGCACGGTTTCGGGAACAAGCACCATTGCGCCGGGAGTGGCCATCACGGGGAGCAGGGGCAGC
GGACTGCTCCTCGAGCAGAACCTGACGGGACTGTACCTCGACGGCTACCGGCTGAACTGCACCAACGAGACGCTCAACCTAACCGACTCCTGCGGCGAGCTGAGAGTGGTGG
ACCACAACTACTGGGCGCTCATCCTGATTCTGTTCCCCATCCTGACCCTCTTCGGCAACATCCTGGTCATCCTGTCCGTGTGCCGCGAGCGCTCCCTGCAGACGGTCACGAA
TTATTTTATAGTCTCGTTGGCCATCGCCGATCTCCTGGTCGCCGTGGTCGTGATGCCATTTGCTGTCTACTTTCTGGAGCCAGTGGAGCCCACAACATGTATAATGGGGTCA
GGGGTCAGTGGAGCGGCAGAGGGGCAGTGGGGCGGAGGAGCGTAG
(SEQ ID NO: 32)

Start ATG: 1 (Reverse strand: CAT)

MLSPFDWRRGISSSGTGGTMAAQPLSSTAATTAAATGATAATAATAATTSATLSTAAASTSTTAAPSAGATWINHHLAVEADSSQPANGSDAQAGVEGPTMPAGYLPLYEDV
ETAAEDAGYALIDDISEWLLGSVGSEAAVGGPENSTNLAVTGANGTLAWLEALNSTQPAQSNSSAEDGERGRYSLRSFVEQQLAGGGAAGAGDGGDAGIALIDSGEEAALDN
VADAETDYGMLGGFGDAELLQRTATVARETLGNRTAPSTTSYDGGGSGDVGVAGGLAGTAGGGVGGAGGSGGSTFMLLLENFNDYFPNYNGSTVSGTSTIAPGVAITGSRGS
GLLLEQNLTGLYLDGYRLNCTNETLNLTDSCGELRVVDHNYWALILILFPILTLFGNILVILSVCRERSLQTVTNYFIVSLAIADLLVAVVVMPFAVYFLEPVEPTTCIMGS
GVSGAAEGQWGGGA*
(SEQ ID NO: 33)

Name: G-protein coupled receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384638
TTTGTACCGCAGTCAAGCGCAAAAATTCCTCGAAAAACCATAGACCAAGGTGTTTATGAGTTTATTAATACTGTGTAATTATTATTTGTGATATATGTGTGGGAAATTGGTA
ATAGATCAAACATGAAAGTAATGGTATCTGTTTTTCTGTATTAAAATAACATATCTGATAATTTCCTAAAGGATGTTAATGTTAATTTAATAGGTATTCCAAAACTAAGAGT
GAACTGGCAAAATGTAAAATTAGTTGGTATTTCCTTTTGAATAAGCTTGACCTTTAACTTACTTCATAGCCTATGACAACTCATATATTATAAATATAAATAATAATAAATG
AGTGCTAATAATAATTTTTAATCAAAAATTCTCACAAGCTAACCTTGTAATGATGGTACGCCTAACTGATACCCATTATGTTACCATGGCCAGGCATAAATTCCTGTTACTG
GCACCGCATCATCATCATACTGGCATTCGGTCATATTAAAATAATATCACCTATCTATAGGGCTGCGGCTCTTTTGACACGCACACGTGCTGCTCAGCAGACCAAGTGAACT
TATCAATGAACACCCGCCACACCCACACGGTGGCCCAAGCACCTGACCCACCCATGCCCAGTGGCCAGCACCACCCGAACTCGCAGAAGCCAATGAATGGCCGGCAGACCAA
GAAACACCCACTAACATCTCGGGAGATTCTTCGTCTGCCGAGTTCAAGAACAACACACAAGTGGCCGTAATGAAATGTGAATAGCCTTGGCAAACATGAGTCCGGAAAGCCA
ACTTATGATGGCTTCTGAATAGCCAACCGCATAGTACACGGCTAAGAGCGGGGACCAAAAAAGGACCTGCTCTTTGCTCGCTCGTTGGCTTTCTCTTTCGCAGCCTCGGCTC
GGAAAAAGGATGCCGACAAGCAACAATTTTCATTGATAAAAACACCAGCTGCTGGGTCCTGAAACTCAGTCGTACGCGAGTCCTCAGTCAGTCAGTTTCGAGCTCTGGGATG
CAACGGATGCGGATGCCTTAAATCCCCGAGCCGTTTGGCGGTGTTATTGAAAATTGCGTATTCGCCGCGTATTACCAAGTAATAGCTTGTCCAAAGCCACTACGCTGTCAAA
GGACGGGCTTAAAACGCTACTTGCATTTTTACGTACTATACATAGATAAATTTTCCGGGTAATCGGCGGAAATTGAGTTAAAGTTTTGTGTCCGGTTGGCGCCAGTTGGCAA
TTGGTGCTGGCCAACTTTCGGACCCGAACGATGACAAACTGAAAAGTAAACTTTGTAAAGACTCAGCACCTTACAAACAAATTCAAGTGTTTCGCGAAAAAAAATAAAAAAT
TTTGAATTGAAGAGGACACAGTTCGGTGAAAAAAAAATATTTGCAATGCACAATGGAAAATCAGCAAAGTGACCTCCAAGTTTACGAAAACAGAAACTAACTCAATAACCAT
CCCTGCAGGTGAGCAGAGGCAACATTTTTCTTTAAAACATTTTTTACCATATTTCCGTAAACGGTTTGCGTCCTTTGGCAAATTTATGGTTGTCTTGGAATAAAATGAACT
TTATTTACTTGGCAACAAAACGAGTGTTTATTCTTGTCGCATAGGGAATCAAGCAGCCAACATAAGGACTATATTTTGTTTAATTCGATTTAATGTGAGTTGCAGAATTAAA
TTTCTTTTTAATTCAACACGCATCTGTACACCATTTGTAAAATTACAAGCTAAAGCTCAGCAAACACTAAATTTTAATCAGTGAAACGAACTGTATTCTGTCAATTTTAATA
TCAAAAGCATATACATTGCTCTTAAAAATATTAAACTGCGTTAGTAACGTAAACAGAAAATGCTATGAAATTTGTTATAATCAAAATTCACTTGAAAGAGATTTCAACAAAA
TTACCAATCTCGTGTTCTAAGGAATTAATATAATAATTATAATAATATCATAGCTTGTTCCCTGTCACCTAATATATACCATTTAAATAAAAGTAATGCCTAGGTAATTGTT
TGTGAGCGATAAGCGCAGATTTGTCTGCATAATTATGTGTTAATTTGATTGAGGTCTACTGAAGAGGGCCTTATTTTTATAAGCCAAAAGAACAAATTCAAACAAAACACTT
GAGGGCGGCCTTAGCTGGCAATAAGCGGCTTTTGCCTAATGGAAAATAAGTGGAATGCCTAGGAATTATTGGCCAAGTGGACGGGCTGCTCTCTTTCTGGCAATCCTGAGTC
CTTTCGACTGATTCGTTTTTGGCTCCTTTGGCCCCATCTCCGCACGCTTTGTTCGGCTTGACACACCTGAGTGGCGATGTTCCACACAGGCACTTTCCCAAACAAATGGAAC
CCGGCATCTCGCATTTGGTTGCCCGAACGTAAGTCCCATGAAGTGGTCTTCTCCAGTGGACAGCCAGTCAGCCGCATTTCTTGGGGGCGCTCTTCTCGAGCGCTTGGCCTCT
GAGATGGCCTCTTTGGATGTCTTGGAGATATGCAGCTCATCCTGGCCACCCACTTCGCCGAGCCATTCCATTTCATTTCCCTGGCCAACATACCCGTATGAAACGATGAGTG
CTCTTCACAGTTTTTGCCTACATTCCGGCCACTGCACCGACTGATGGTTGATGGAGACCCATTCTCCGGTTAGTGTTCGTCTTGGTCGTCTTGGTCTGGGTCTCGGTCCTGG
TCACAATGGGAATTCGCAGCTTAAAGGCGCCGTGGCTTTATTCAGGCCAGATTGGATACTTGACATGTTCTAATGCGGCGCAGATCTCAGTTTCTGCAACGTGCTCTCCCTA
TCTCAGTTACATGAATTAGCATTCATCGCCACTGCACTGAGCAAAAAATAGGACCACAAATCGCTCTAACTGCTGTAAATTGATATCCATTCAATTATTTATTTAAGTATT
ATAACACATTGATAGAACATTAACATGTCATATATTTTTGAACACAATTCAGCATTACACTTTAATAAAGTCGAATTTGAATTTATCCAACTGACGCGTTGGGTGATATCAT
AAATTCAATTTCGAGAATGACAGGTCATTAAAATGTTGTCACCTACACTGTCAAATATTTAAGTCAAAATATTAAGTTATAAAATATCAAAAAGCTTTGATAGCTTTAGAGT
ATGATTTACGTTTAAAAGGATTGTACACGAATACAATTTTGTTTTTTAAGCAATCAGCGGCGTGTTTCTGTTTTGCCATCATAAATCGCAGAGTTTCATAATCCCTCTAAA
TTAAAATGTCCTTTGGGATTTTTCTGAAATTTTTTTCAAACGCTTCCCCCTTTTTGTTTCCTGCATGTTTAGGCCGTTAAGAGCCTGTTGATGATGGCGGCATGGTAGAACT
TTCTGTTTTGTCACAGCCTTGGGCTGCTGGCTTTACATTTGGTGGAAAAACGAGACCAGCCGAGCACGTGCGGTCCGCGGGGTTGGGGGGGTTGTGGAAGCTAAGGACGTG
GCAGGACGAAAGCGCTGAGTGCAGATGAGGACTGCAATAAACTGCCATACATAAGTGAAAATTGTGAAATGAAAACTTGCGCGAGCATAAATCAATACACCCAAGGGGGGTC
ATGGGCTCCCATGGCAGGCAACAAAAAAATATATATGTATATATATCAGCTGGGAAATGTGCTCCCTACTCAGGTTTAGCGACTTGTATTTCATGTTCCCAAATCAGCTGGT
GTTATCCTTTGATTGCCGGGCACACGACAAAAGTGGTGGACCAGGACACATGTGATGATTGCTTGCGTGGGTCGCAGGAATGAGCCAACTCAGATTGGTAATATTGTGTATG
AACGTGATTTCGTGACTTTGCAATGTTTCTATTTAACAAAAACTCTTGTTAAAGAAATACACACAAAACATCACTTATAACCCTTTTTGGATATCTATGAGGATTTAAGGTT
TGTAAACTAAGGCAATGGTATACCTTAAAAAGTCTATAGTAAATCGCATATATTACATCAATTTTAACGTTCATGCTGATCTTCCAGTTCCAAAATATATATTTTGAAATAG
CTTACATTTAATTAAGTGCACTTGCTATAAAAGTTTTGAGCTTGCCTGAAGTTTATTACATAAAAAATTCGCAGAAAGGCGGCTAAAGAATGAAACTAATTTAAGCCCAAAT
GATAAACTATTTGCCAAGCATGTTTCGTTTGAATTATGCTCCCGTTGCTCTGCCTACCAAAAATGGCAAAAACTCGCAATATGATTATTGTGGCAGCTTCTTCCGGTGAATT
AAAAATACAAAAAACCAGCTCAAATTGAAACTTAATTTTAATAAATTGTTTTCATGCTCTTTTTCGCTGCCGTTGTAGCGTTTTCCCGCCCATTTCACCCGAGTGGCTGGAT
AATTGTATCCGCGCAATCATCCCGCGGCTTCCATTGCCGCAGGATAAAAACCAAACTGCAACGAACAGAACAGAACTGAGCACTTTAAGCCATTTCCACTCGTAATTAAATT
TTGAGAGCGGATGAAATTGTGGCGGCTAAATAAAAGGAACAAAATGCATCTGTGCATACAGCAAAAAAAAAACAGCAAAAATAAATAACAAAATGCAGAAGAGCAAAAAATC
GCGCATATCCGCTTTGTAAAAATACGAGTCTAAGTGATGGGTTCCGTTTTTGTTTAGTAGTGCTCACTAGTGAACGGCCTTGTCGCATTTGGGGCTGATAGTGGATGTAGTG
TGGGGTGTATTGTGGTGTGCATGGGGGTGTTTGAATCGCAGAAGTCGGCCCAGCCCGTTGGCCGTCAAGAGGACCTGGCCTAATCGAATGAATGTAGAATGGGCACAAACAA
GTCACGCCTGCCTGCGCATTTCATCAAAGAACGGCGAACAAGAGGCTAACAGCTCTTCGTTTCTCTCTCTCTTTCTCTGCTCAGCTGGGTCCTTTGAAGATCCTGAATGACA
CTCAATTGGCCTAATCGGTTTCTGAAAAGCATAGTATGAATTTCATACGGTACGCTAATTTCGAATATAACAAAAAGCAAAAACGCTTTTGATATCTAAATTGTATTTGCCA
TTGAAAAGTTTATACCTATATGTATAATATGGAACCTTTCTAAATATAGTAATTGATACGAAATTGCAAAAATGATGGATGAACATCTCAACATTCAACAGGTTGCCCAGTT
ACTTAATCAATCCATCGGTGTTGAAAGTGAGCACTTTATACAAAAATACTCATGAGTGGTTCTAAAGCGGATAAGAGGATTTATCTTGAGCTCAAAAACTGATGCAGTGATG
CAATATTCAAGATGTGTGCCTTAGTATTCAACATTTTTGCCAGATTTATCGAAACGTAAATAAATAAAAATATACATGAAATTTAATAACATTGCGCAGCTTTTTGGTGAAA
GCAGTTTTAATTTAATAAAGCAATCGTTGCGTTGAAACTAATATTATTTTAAATTATACAGGATACAAAAAATAATATACATATACATTTAAATTTTGAATTTAAAAAGGGG

TGTGACAGATTTACTGAATAATAATAACTTTAAACATGCCACCTAACAAAAGCGAAAACAATTGTTTAATTACATTTGTGCAACTTGTGGCATTTAGTAAAAGCGAATATCT
AAGCAAATAAGTCGTGTAATTTGATATGCAATATAGCAACCAAACAAAAATGCGTGTTATTAAAATCATGATTACTGCCATACGTATAGATGTTTCATGTCCTGTCCCTTGT
TGATTTCATGCTGATAAAGGACATACACTTTTAAATCTTGCTAACTTGGTACTTTTAACGAGCATTCGAAGGTATCCTGTGCCCATTTCAAAAAAGCTGTAAGCTCTTTGGC
AGCCAAAAAAGCTGACAAATTTGTGCCGGAAAGGAAACAACTTCCTATTTGCATGATTTATTGCCAATGGAATAACAGGTTGGCTTCAGTTGCGTATGTCCCTGCTATTCGA
TTTTCAACGAAATTTCGGTTGGCATTCAGTCGGAATGAAAGTGAAAACAAAAAGTGCTACTACTACCCTTTTGTGAAAGCCCCCGGAGAAAAGACAACTTTCAATTGGCCAA
TTAGTGCTATTTAAATGGCAAGTTACGCCGTTTTCAAAGGGCCAACTATTTTGTTTTTGAAGTTCCGCGAGTTGTAAATGTAATAAGCAAATTTCGCAGTTCTCCATAGGAC
ACCTTTACACTGAGTTCCTTTCCAGTTGCTCAAGTTTTGCTCCTTGGCAGCTGCACTTGTTGTTCCTGTTTGCCAGAAAGTTAGTTTCTGGCCAAGAATTTGTCTTTGGCCA
AGTACTGCTCACGCTCGGTTTCTTTGGGCAAAAGTTGGTATTTGAGTTTCGCAGTGGAAAAGCACTTGCGGCAGATCTTCAAATTCTGCCAAACCCGGCGGGCTCACGTGGC
CGCCTTCTAGTTATTTAATAAGTAATTAAGTTCCCTAATTAAATACATTCTCGCAGAGATTTACATAATAAGTAATCTACGCCGAACGTCTTCCTCCATTGCGGCGTTAATT
AAATTTGCAGACCGATTGCATTTCGGAAACTTGATTTGGTTTGCGTCTGGCTGCCAGTCGTTAACAAATCATGATTGCTGAAATGGTCAACTTTGGGCATCGCACTTCGTAT
ACACTTGAGATTAAAAATTCTTTAATAAGGGGAAATAACACTTATTTATAGAAGTATTTTTAAAATATATACAATTAATCTAGTCAGCGCTTTTGACGTAATTTATATTGAG
TTCGAACAACTCATTTGACAACATCAAAGAATCCAATATTCAGCAATTTTAAACATACTCACATGCGAAGGGGGAATCGAATTTCCAATAATTTAAAAAGGAATCGGAATTT
TACATTTTCTCGCGCTTTTGCGGAGAAGCCGTGATATGGTGACTGTATTTATGAAATTTGCTTTTACGCCACTTCCATGAACTGAAAATTGGTTTTCACTGCCTGTTTTTGG
TCGATATAATAATGCTGAGATATATATATGATATGTAATGTCTTTCTATATATTGTAGCATTTACGAAGGGTTATAGCTTGATATAAAAAAAGAAAGAAACTTTGAACAAAC
TCAAGTATATAAAAAGGTTTCAATAAAATATATAAAGTCCATTAGCTGTAGCTTACAGGAGCATAAGTTTTCTTGTTTAAGACGAAGCTTTATCGAACTTTATCTTATAGAT
ACATATATCTATTACTAGGATAAGTAAGTAATTACAATTATTTATGCACGCTGCATTGACATATGCCAAACACTTAAAGTATTCCCACCCAGCCGCACTTGCTACATTTATC
AGCAATTGAATTAATTTCAGGTTGCGGTTCAGCTGAGTTTGGTTTGTAAAATACATTATAATGTTTAATAATCAGCAGAGAGCAGGAAAAATATGTATTAAGTTCGGCATTG
TGTTAAACTTTAAGCGAATAGTTAACACAGTGTTATTGGTTACCCGCTGCTTTTAATAACTTTCCCTTTCGATGCCCACACAACCTGAGTTATTTTCCTTAATGGTCTGAAA
AAATGAAGAAAAAAAAGTTTGGCTCAAGTTTATTGGTTTCGCGCTGCTTTCATCAATTGCTGCACGCAAATGCTTCTCAAAAAAAAAAAAAAAAAAAACTAACTCTTCATTTA
TGGCATTTCTTAGTTGCGAATAAATTCCGATTTAAGCTGGTTGAAAAAAACCCATCCTTTCAACGGTGACTAAACATATTTTAAGCAGACCGCTCAAACTCAAACCGTGGGT
GGTTAATTAAATAAGAGCATCGATTAAGTCAACTGTCTTTTTATCTGTCCCACTTTTAGGGTCACCCGGTGGACAGCTAATTGCAGCTAACCCCATTGGCATTTAGCAAACG
GAAAATGCATCGTGACCGCTGCAACCACATTGCACCACTACCACCAGCAGCATCAGCGCCACCACTTCCACCTCCAGCAGCAGCAAGGGAATCCGCACCGCTGAAACTGAAT
CTGAAGGCATCCCAGTGGCGGCGACACACGGAATCCTTGCCAAATGCGGCCAAATGCGAGGCGCTGGCAAGTAAATAAGCGTAATCCTTTGCGGGCACGTACCACCACTACT
ACCTCCTGTTCGGAGTAGCTCGTAAATATTTGGCCAAGTCAGTCATGGAGGCCATCATCATGACGACCCTGCCCAACCTGACAACGGATGCAGGTGACAGCAGCTTCTGGCT
AACCGGTGCCCTCTCGCTCTCCGAGATGTTGGCCAACTCGAGCCACAGCCATTCCACTGGGAGCACAACCTCGACGGCTGGGAGTTCGGCCACCGAATCATCCGCTGTGAAT
GTAGGCAAGGACCACGACAAGCACGTGAATGACAGCGTTTCCACGGGTCTGAGGTAAGTGCATGTGCCTAGGCTTTTCCATTTCCATGGCAACACACACACACAGATTCCATTA
AGCTCATGACTGCGGGCTCAACAAGTCATCGCGATAAGCTCAGCTTCATTTTCCAGGAGATGCATTCAACTAATGTGCTTAGAATTCTACTATTGTACATAACTAACATATA
GGGAATATAATCAGATTACTTAGAACCTCAGAACAGTCAATACAATCCACTTAATATAAATGAATTAAAAGTATCTATCGAAACTGAAAAGAAGAGGATCATTAGCTTGGGT
ACAACATTTATAAGAATTTTGGCTTATGAAGTTGCAAAAAGAATTCTTAAAAAATTCAATACAATGTCAAATAGTTCTGAAACCATAAATTATTAACAATTAAATATTAAAT
TATTTATATTGATTTAATTCATACATTTAAACATAAGTTACACTACTGCAAGGAGGAAAGTATCAGATGACTTGAGCTATACTTTCATAAACTTAGCTTAGAGCAATCTAAA
CCAAATTGAAAACTCATTGAAACCCATTTTAATTGAATAAATTGAAACATTTACTTTTCCCATTTCCCCAGCAATTACAGTAATTACCCCAGCTACATTCACTACAGGGACA
AGTACGACCTGAGCTACATTGCCAAGGTGAATCCCTTTTGGCTACAGTTCGAGCCACCTAAGTCGAGCACCTTCCTGATTATGGCCGCCCTGTATTGCCTGATTTCGGTGGT
GGGATGTGTGGGCAATGCGTTCGTCATCTTTATGTTTGCCAACCGCAAATCTCTGCGGACTCCTGCCAATATCCTGGTGATGAATCTGGCCATCTGCGACTTTCTGATGCTC
ATCAAATGCCCGATTGCCATTTACAATAACATCAAAGAGGGTCCAGCTTTAGGAGACATTGGTAAGCTTAAAACTTGGGGAAATAACTTACTAAATATCTATATAAATGAAA
TTCATTCGGCTTGCATTTATGTCAAAAATCGATCAACAAAAGCAATCTACATAATAAATAAAATTAAATTGGCCCAAAATGCTTGGACTGATTTCAGCTGTTTCTTGCAGAA
ATTCATTTGTGGGATTCTCTTTTCCCGAAAGCAATTAATAGAATTTAAATGAAACAATAAAAGTTCAATTTGTAGCCTACACATCTTGGTTCATGAAAAACATATTCCAAAA
TGTAATGGCTATTAAAATTTCACCGAGCCTTTTGTTTGCCTTGAAACGGGTGGTTCATTAAAAATCCTACTGACTTTCCGTCCCACAGCCTGTCGCCTCTATGGATTTGTGG
GTGGCCTAAGTGGCACCTGTGCCATCGGTACCCTCACCGCCATCGCTTTGGATCGGTACAATGTGGTGGTGCATCCACTGCAACCGCTGAGACGCTGCTCCCGCCTGCGATC
CTACCTGATCATCCTGCTGATCTGGTGCTACAGCTTCCTGTTCGCTGTGATGCCGGCCCTGGATATCGGACTATCTGTCTATGTGCCAGAGGGCTTCCTCACCACCTGCAGC
TTCGATTATCTGAACAAGGAGATGCCAGCTCGCATTTTCATGGCACTGTTCTTTGTGGCTGCCTACTGCATTCCACTGACCTCCATTGTGTACTCCTACTTCTATATACTGA
AGGTGGTCTTCACAGCGAGTCGAATACAATCGAACAAGGATAAGGCCAAGACGGAGCAGAAGTTGGCCTTCATTGTGGCGGCTATCATTGGTTTGTGGTTCCTGGCCTGGTC
ACCGTACGCCATCGTTGCCATGATGGGTGTCTTCGGCCTGGAGAGGCACATAACGCCACTGGGATCGATGATACCCGCGCTCTTCTGCAAGACAGCTGCCTGCGTGGATCCC
TACTTGTATGCGGCCACCCATCCGCGGTTTCGTGTCGAGGTGCGAATGCTCTTCTACGGACGCGGAGTGCTGCGAAGGGTCTCCACGACCAGATCATCGTACATGACCCGAT
CACGATCGTCGTTCACCCACCGATTGAGGACGAGCACCACCGGTGAGGGCGGAATGGGTGATCACCGCATGGAAAACTATCTGATGAACAACAATCTGATGATGGTGCCCGA
GGAGACGGAGGAGAACGAGGAGATCGTTGTGGTGGCCGAGATCAACAATTCGATTAGCAGCGTCATGGAACAGAGTAAGTTCTGAAGGCTGACGATTATGCCAGTTAAGAGA
ACGGAAATTCTGTCAATCTGTTTCGAACCAACTAACATAAACATTTTTAATCAGTTTTAAGTTTTACAGACGTGTTTTATATAATTTGAAAATTAGTCAAACCTATGATCAC
GATGATACTGAAAACTTTGCTTGGAATATCTGAAACTCGCCTGTTTAATTTCCACTTTTTCCAATTGAAAATTTCAATTTGATTTTTGCCCAGGTAATTTATAAAAAAACCT
TTCCTATCTCGAAAAGTAAACTGTTCGCTGTTAATTGAAAATGTTTGTGCACTGCACTGCAAGGCATATCAAAATGGAAGCTATCAATAATCGTAAATTGAATGGATCGGAC
TCAAGTTTCATTGTTGAACTAATCTAATCTAATTTGTATGTAGAAATAATAGACCGAACTGAAATCAGCAATGGCAGCAAAGTAATTTCGCTCGGCAATGATATATGCCAAT
GGCATTGTTCATCTGAGTGAATGCGTTGTCATTGATGGACCTACAATTGAAATGAAAATAATAATAATAATAAATAGAAATGAAGCAGGGAAATCCGCCGAGCAGTGGGAAA
TTGGCAAAGGCACAACGACGACACAGTGGTGATAGTAAACATATTTAATGAGCTTATGTGATGTTTAGACAGGGGGAGTAGGCGATCAGAGCAGGGAATGGATTAGTTAGCC
TATATACATATATATATACACAGGCATATGTACATATGTATAAGACATCGGAATAGCAGGCCAATTAGAGAGGCAGGGCCGAAATTAGTACAAAGAAAGAGGGTGGCGTGGG
TAGGGGCCGAAGTGGAAATTCAATAGTCAGATTATACATCATTTCAGGGAAAGTCACTCGGACCCAAGGTCAAGCCGTGGTATTGGATGTTAACAAGCGGATGTGCAGATTG
TTTGCATTGTAGTTTAATATTGGAATAACTCACAATCAAAGCAGCCTGCGGACATCTTGTATTTCTGAAATGGAGTC
(SEQ ID NO: 34)

Exon: 1001..1464
Exon: 7788..8022
Exon: 8083..8341
Exon: 8920..9245
Exon: 9609..10501
Start ATG: 7845

Transcript No. : CT17820
CTGGGATGCAACGGATGCGGATGCCTTAAATCCCCGAGCCGTTTGGCGGTGTTATTGAAAATTGCGTATTCGCCGCGTATTACCAAGTAATAGCTTGTCCAAAGCCACTACG
CTGTCAAAGGACGGGCTTAAAACGCTACTTGCATTTTTACGTACTATACATAGATAAATTTTCCGGGTAATCGGCGGAAATTGAGTTAAAGTTTTGTGTCCGGTTGGCGCCA
GTTGGCAATTGGTGCTGGCCAACTTTCGGACCCGAACGATGACAAACTGAAAAGTAAACTTTGTAAAGACTCAGCACCTTACAAACAAATTCAAGTGTTTCGCGAAAAAAAA
TAAAAAATTTTGAATTGAAGAGGACACAGTTCGGTGAAAAAAAAATATTTGCAATGCACAATGGAAAATCAGCAAAGTGACCTCCAAGTTTACGAAAACAGAAACTAACTCA
ATAACCATCCCTGCAGGGTCACCCGGTGGACAGCTAATTGCAGCTAACCCCATTGGCATTTAGCAAACGGAAAATGCATCGTGACCGCTGCAACCACATTGCACCACTACCA
CCAGCAGCATCAGCGCCACCACTTCCACCTCCAGCAGCAGCAAGGGAATCCGCACCGCTGAAACTGAATCTGAAGGCATCCCAGTGGCGGCGACACACGGAATCCTTGCCAA
ATGCGGCCAAATGCGAGGCGCTGGCAACTCGTAAATATTTGGCCAAGTCAGTCATGGAGGCCATCATCATGACGACCCTGCCCAACCTGACAACGGATGCAGGTGACAGCAG
CTTCTGGCTAACCGGTGCCCTCTCGCTCTCCGAGATGTTGGCCAACTCGAGCCACAGCCATTCCACTGGGAGCACAACCTCGACGGCTGGGAGTTCGGCCACCGAATCATCC
GCTGTGAATGTAGGCAAGGACCACGACAAGCACGTGAATGACAGCGTTTCCACGGGTCTGAGCAATTACAGTAATTACCCCAGCTACATTCACTACAGGGACAAGTACGACC
TGAGCTACATTGCCAAGGTGAATCCCTTTTGGCTACAGTTCGAGCCACCTAAGTCGAGCACCTTCCTGATTATGGCCGCCCTGTATTGCCTGATTTCGGTGGTGGGATGTGT
GGGCAATGCGTTCGTCATCTTTATGTTTGCCAACCGCAAATCTCTGCGGACTCCTGCCAATATCCTGGTGATGAATCTGGCCATCTGCGACTTTCTGATGCTCATCAAATGC
CCGATTGCCATTTACAATAACATCAAAGAGGGTCCAGCTTTAGGAGACATTGCCTGTCGCCTCTATGGATTTGTGGGTGGCCTAAGTGGCACCTGTGCCATCGGTACCCTCA

CCGCCATCGCTTTGGATCGGTACAATGTGGTGGTGCATCCACTGCAACCGCTGAGACGCTGCTCCCGCCTGCGATCCTACCTGATCATCCTGCTGATCTGGTGCTACAGCTT
CCTGTTCGCTGTGATGCCGGCCCTGGATATCGGACTATCTGTCTATGTGCCAGAGGGCTTCCTCACCACCTGCAGCTTCGATTATCTGAACAAGGAGATGCCAGCTCGCATT
TTCATGGCACTGTTCTTTGTGGCTGCCTACTGCATTCCACTGACCTCCATTGTGTACTCCTACTTCTATATACTGAAGGTGGTCTTCACAGCGAGTCGAATACAATCGAACA
AGGATAAGGCCAAGACGGAGCAGAAGTTGGCCTTCATTGTGGCGGCTATCATTGGTTTGTGGTTCCTGGCCTGGTCACCGTACGCCATCGTTGCCATGATGGGTGTCTTCGG
CCTGGAGAGGCACATAACGCCACTGGGATCGATGATACCCGCGCTCTTCTGCAAGACAGCTGCCTGCGTGGATCCCTACTTGTATGCGGCCACCCATCCGCGGTTTCGTGTC
GAGGTGCGAATGCTCTTCTACGGACGCGGAGTGCTGCGAAGGGTCTCCACGACCAGATCATCGTACATGACCCGATCACGATCGTCGTTCACCCACCGATTGAGGACGAGCA
CCACCGGTGAGGGCGGAATGGGTGATCACCGCATGGAAAACTATCTGATGAACAACAATCTGATGATGGTGCCCGAGGAGACGGAGGAGAACGAGGAGATCGTTGTGGTGGC
CGAGATCAACAATTCGATTAGCAGCGTCATGGAACAGAGTAAGTTCTGA
(SEQ ID NO: 35)

Start ATG: 522

MHRDRCNHIAPLPPAASAPPLPPPAAARESAPLKLNLKASQWRRHTESLPNAAKCEALATRKYLAKSVMEAIIMTTLPNLTTDAGDSSFWLTGALSLSEMLANSSHSHSTGS
TTSTAGSSATESSAVNVGKDHDKHVNDSVSTGLSNYSNYPSYIHYRDKYDLSYIAKVNPFWLQFEPPKSSTFLIMAALYCLISVVGCVGNAFVIFMFANRKSLRTPANILVM
NLAICDFLMLIKCPIAIYNNIKEGPALGDIACRLYGFVGGLSGTCAIGTLTAIALDRYNVVVHPLQPLRRCSRLRSYLIILLIWCYSFLFAVMPALDIGLSVYVPEGFLTTC
SFDYLNKEMPARIFMALFFVAAYCIPLTSIVYSYFYILKVVFTASRIQSNKDKAKTEQKLAFIVAAIIGLWFLAWSPYAIVAMMGVFGLERHITPLGSMIPALFCKTAACVD
PYLYAATHPRFRVEVRMLFYGRGVLRRVSTTRSSYMTRSRSSFTHRLRTSTTGEGGMGDHRMENYLMNNNLMMVPEETEENEEIVVVAEINNSISSVMEQSKF*
(SEQ ID NO: 36)

Name: Rhodopsin 3 and 4-like (Rhodopsin 7)
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384543
TATGCGCAACTCGACAAACATCTTCCTAACCAACCTGAGCATTGCGGACCTGCTGGTGCTGCTCGTCTGCACGCCCACCGTCCTCGTGGAGGTCAACACTCGCCCGGAGACC
TGGGTACTTGGCCACGAGATGTGTGAGTGGCTTCAGCTCCAGTGCTCCTGGTCTATTTGTTCTCCCCAACCTAATTAACTGCTGTTCTATCCCAGGCAAGGCTGTGCCGTTC
GTGGAGCTCACGGTGGCCCACGCCAGTGTGCTGACCATCTTGGCCATCTCGTTCGAGCGCTATTACGCAATCTGCGAGCCATTAAAGGCCGGCTACGTCTGCACCAAGGGGC
GGGCCATCCTCATCTGCGTCCTGGCCTGGGGCATTGCTGCGCTCTTTACGAGGTAAGCGATTAGGCTCGTACTTGATTTCTTAAATTTATTTGTAAACTAGCATTATACTT
TAAATATCTTAATTAGTTTGGTCGCTTTTCATCGCATGCTTTTTGGTGTTTGTAGTCCTACATTTCGTATGTACATATGGTATTATCCTGTTTAGATCTACTTCGTTTGTGG
CCTATATCCGTTTATTCATTTCGCCCAGTCAAGCATTAAGTTCCTCGCTCATTAATTATAAGTGTCAATCGCCGTCATTTTAATGATTTGTGTTTTGTTATTGGGTTTTGCA
CATTTTTCCCCGTTTGTTTATTTCCATTATGACGTGGAATTTCTGGTCTTGGTGTTTGGCCTTTTGTGCGCCATCTACGCTGGCGGCGTGTCCATTTGTCCATTTACATG
CTAAACTGCGACTGTCGCCTATCTATTCCATTTCCACTCCTGTCCTACCACTTCTCCTGCTTCACTCGCTTCATCGTTGCCATTGCCATTTCCATTCCCAATCCCATTCCCA
TTGACATCCGTCGCCTCAAGTTGCGGACGGCTTTTGTCCAAGTTCCGGAAGCTTGCCGTTATCAGCGCCCTGTCATCTTTAATAAGCACCGATGTCCGCTATCATGGATTTC
ATTTCATTGACACTTTATTATGGTGCGCTCCGTCCAAGGACGAGCGGCGACCCTTTTTAATTTTGCAAATCTTGCCGCGTCCATAACAATGAAATAAACAGCACATTGTCGA
GCTGGCCACCGACCCCTCAGCACTCCGGGAGTCATGGGGTCATCTATGGTTGGCAGCTGGGTCCGCAGGACACCCGCCTACGGGCCGAGACCCATTAATTACGCCACACAAG
AAGCCCGAGGCCAAAAAAAAAAAAAAAAAAAACGGGCAAACATTTTTAAACACATTCAAACAGACGCGCCACACGACGCCCATGGGCGCAAAAGGAGGCGTGTCAAATAGCTG
ACGCACGTACCAAAACGCCCGGATCATCTTTAGTGGGAAAAAAAACACGGCCATCGCCTGTTGACTGTCATGGCATTTCGTTTTCGAGGTCTAGCCAGGGAGGATTTGGGCT
GGGGCTGTGTTTAATCAGCGGCTCGGATAACTCAAAATAATCAAATATTTAAACTTGTCTGTTTGCTGGATATATTTACAAAGTCAACAGAGTATATCCCACCATAAGTTCG
CCATCTACTCACTTATCCCCTTGTAAGTGAAATGTAATCACATCAATGAAATTTGGAATATTTTGAGCTTTAGTTAAATGTATTTTAAGATGTATCGATATAACTCATACAG
ATACTTTTTTAAAATGAGTGGTTTCATTTTTATTAAATTTCTGTTATTATAATAATTTTAAATGCAATGTTGATATTATATTTCAAATTGGCTTAACGCCTACAGATTAGAA
TACACTAAAGATACACGGTATACATTTTTTTTCTAGACAGAAATATACTTTAATTCAACTAAATTAATATATATATATGTACCTCTCCCATCTTTTAGCTATTAAATAGTCGAG
GAACTCGACTATAGTATTATATCCGGTTTTGGTTTAAAAACCAATCATTGTTTTTTTTTAACGACACAGTACTATTTAATATGGGCTTCAACATATATTTTAAATGATGAAC
CCAAGAAATGAAAAAACCATAAAAACAAGCGGTCATGAAAATCGCGTCCTGGCCAGACTGGTACATTGGAAACTTGAACTTCATCCCAGACTGGGTTAACCTTTTGCCCATC
GCCATTCTCTCATCATCACAAACAGCTTCTCTTCCTGCCTGCTGAAACCTCTTCTTTTTCCCGCCACTCTTGGCCTTTTCCTCCAAAATCCAACCGTATCCATTATCCATTA
TCCACGCCCCCGGTAACCCCGCCACCTGATTACGTTTCTTGTTTGCTTTCTCTCCCCAATCTGCAATATCCCCCGAGCAGCCCCATTCTGTGGGTGGCCGAGTACAAGCTGG
CCGAGTACATTGATGGATCGTCGGTGGCCGTGTGCCTGACCCAGGCCATCAGCGACTGGACGCTGGCCTTCTTCCTGATGACCATCTCGGTGTTCTTCGTGGTGCCGTTCGT
GACCCTGGTGGTGCTGTACGGCATCATCGCCCGGAATCTGGTCTCCAACAGGGCGGCCATGCTGCGCGCCCGTCCCACGAAGCCGGAGCTTAGTCTGAAGGCCCGCAAGCAG
GTGGTCCTAATGCTCGGCGCCGTTGTGCTGTCCTTCTTCGTCTGCCTGCTGCCCTTCCGCGTCCTCACCCTGTGGATCATCCTCAGCACGGACCAAACACTGCACGATCTGG
GATTGGTGCGCTACTACAGCTTGCTGTACTTCTGTAGGATCATGTTGTACCTCAACTCGGCCATGAATCCGATTCTTTACAACCTGATGTCGACCAAGTTTCGAAGGGGTTT
CAAGCGGCTTTGTCAGGACGCAGGGCGATTACTGCTGGAATTGGTGACATTGGGAAGAAGGAAGGAAGACTCTTCTCGCGGGCGCAGAGGTACCTTGTCACTGGGCATGGGC
ACCAATACGAACACGAATACCAACTCCTCAAATGCCACAGGAGCCACGAGCTCCAGCATTCTCTCGAGAAGCTCCAATCGCAGGTGCAGCGAGGATATCAGTCGCACCCGCC
TTAAGATCGAGATGCAGATGCCATGTGGCAGCGACCTGGAGGCCATGGCCATGCTGCAGCATTCCACCTTGGGCAAGGGAATTGCCAGGCGAGTGAGTGATAGCCGTCTCAT
GCCTCTGAGAAATCATCAACCACGACGCCACAAGCCGCAAATAAGTTTCGACGAGGAGTCACTGGAGGAGAATAAGCGTAGTGAAGCGAAAATCCCAACAAAATGTCGAGAA
AAGCTTCCTGGAATAGCAAGAGAAATAGTAAATCTGACCGAAAATACCCTCTAAGCGAGATGGGCGTGAAAAATCGCCTTTAACCCGTTTACACACTCGCTTCACTTGGCTG
GAATGTGCCGATGACGCTGATCCCGATCCCGATTCCTATTCCCGCTCCTGGTAAACACAAATCGCTTCCTGAACGCTTTTGGCAGCCATAATCGATTTCAGTCACCCGCTGA
CAGCTGCCGAGAATCCCGAATCGTAGTTGGGCTGCCGCCTTGGACATTTCCGTACCAAAAACAATCAATTTTTCGGGCTATTAAACTCGAATGAATGTGCCAAAGATAAATA
ACAGCAACGACAGCCGAACAGCACTAAAACTACCACTACTGGTCAGGGTTAGGAGTCCAGATATCGACTGGCAGTAAGGGGTAAGCCATTAGGTGGCCTTCGAGGAGCGGAG
CTGGCAAATGGAAATGGCTTAATATGCGTTTGGGCGAACCATCAGGTCGGCGGATAAACATGATTTAGGTGTATTGAAATAGTTTCCGTACGCCGGACCGTACAAAATGGCC
GCCATTACACGGGCGGCCATCACCGGAAGTCGAAAAAACTGTTGCCTACTTGCCGGCAGCTCAAATAAAAAGCTTAGCTCAATATGCACACAATGTCCACATGTAAACACAG
CTGTACATACACATACATACCCCCACACGCACACATAGCCAAATATTAATATTTATATTTGTACGAACATATTTATATACACATATTTACATATATATATAGTCTGCTTATAAA
TATGTACAAGGAACCACACGCACACACTCACACACACACGAGAGTTTGTAAGTAAATAAATTTTTTATGACAACTCATGGCGGGACTTTTCCACATTACCCCATCCAATCGG
AATCTGTAATTTATATAATTCTAAGGAAATCGGACAAATGTTCGTCCTCTGTAAATATACGTATTCGTACCATGCGCGTAGCTTTTTAATTTGTTTTCGTAGCCCTATGACG
TAGACTCCCACTTTTGTAGAGTAAAAAGTAATAAGTAGCTTGAACCCACACCATCCTGCGAGTATGATGAGTATGAGCAGACGCTGCCGTTTCCGTTTCCGCTGTGTGATTA
AGTTGAGTGCTTCGCCCCCGAAACTTTTCCAGTAGTTAGCCGTTGAGTAGCTTTTAAATATTTGTTTGTTTCCTGCCCGCAGCGCGATAAAACTTTGCAGACAGCGCTAGTC
ACTCTAGTCCCCGCTAAAACTGATAATATGATTCTTGCAGTCCCATCATCGCCATATCCACCTACAGCGTGGAACCCTACGGGGACGGAACCGATGCTCCCGTTTGCACCAC
CGCCGCCGATGGTTTCTGGTCGATCTTCTACTTCGTGGGCTGCATCACGGTGTTTTTCTTCCTGCCCTTCGGCATCCTGGTTCTTCTATACGCGGCCATCGCTTACAAGCTG
CTCCGTCCCAACAACGCCTTCCACCGACCAACCTCCCCGCAGCCACAGCAGCCGTCCGGCGGAGCCACCAGTGGCTCCTCACAGGTGCCCAGCACCAAGGGTAACAGCCATC
AACAAAGCAACGGGATGAGGAAGCATCGCAAGCAGGTAATCTTCATGCTGGTGGCCGTGGTGTCTAGCTTTTTTGTTTGCCTTCTACCCTTCCGGGCCTTCACCCTCTGGGT
GATCCTGGCCAGCGCCGAGGATGTCGAGGGTCTGGGCATTGCCGGCTACTACAACCTGCTGTACTTCTCGCGCTTCATGCTCTACCTAAACTCAGCCATGAACCCGATCCTC
TACAACCTGATGTCCTCCAAATTTCGCAGCGGCTTCTGGCGGCTGCTGCTCACTTGTCTGGGCCAACGGCCACATCACCACCATCGCCACCACTACCATCAGAGGCAGCATC
CAACGGCAGGCGGAAGTGGGCGCAATGCGTCCACGCGACAGGAACAGGATGCCGAGGAAGGAGCTGCGCTGGCGGGAACGACCAGCGCCCGACATCCACGTCGCACACTCCG
CCGCGAGGCCACCTTCTTGATCAACTCCATATCCACCTCCTCGGGTACGGATCGCACCACATCATCATCGGCGTGGCGCAGCAACAGTCTGTCCATTTCCGGTCTGAGCGAA
CGGGAGCGCGGCATACTGGGAGCCGCTATCATCGGCACAACGGCTGCCACCGTTACAACCGCCTGTCTGCAGGAGCGACGCGCCAGCAAGATCTGAATCCCATCCTAAATAT
GCATCCTTCGAGTCCCTGTCCACTGCCAGCAACCGTTAAAGTATTTAGTTGAATTCCAACAGCAAATAGTCATAGTTTTAGAAAAATTTGTTGTAAAGACTGAAGTACCTC
AGAAGAAGTCAATAAACGCAAGAGAAGATGTTTTTGAAAATCGAAACTAATTTTCTTCGGGCTTCGAAAAATTGAAGTTTGCAAGATTCGAAATAGCTTTAACTTGATTTTT
AGAACAGAAACGAGTATGAACTTTTCATACGATCATTTTTATCGCACCTATTGTGATTTAATGGGAGATAGGAGGAGACCAATTGGGATATGCAAAAATAACGCAACTATTA

TTATCTTTGTATGTACGATTAATTTTAATCTCTAAAGAGTCAAAGTACTTCAATAATTTTACAAATCTGTCGGAGGGAATGTAAGCTTGCCCAAATCAAATGGAAGCTACAA
TCGTTGTCAGATGGAGATATTTTATGTTAGAAGGAATATCTCGGATACGAAATATTTAATTTAAAATTCCTTTTCATATTATTCTTTTTAAAAAATAATAAGTCCTGTATGT
ATACAATGCAATACATAAAGGTAGAGTATGTCCTAAAAAGAAAAATATCTTTCCGAATATATGTATATAATTTATAAAATAAGTGGATACATACATTTTTCGTAAGTATCCT
AGAAGATTTGCCAACGTATCAGGTTTACAAAAATTGATTACCTATGTGCAGTTAACTTAACGGCAGTGCACTAACTAATCTGATTTGTGAAATAATGCAGATATAAAACATT
TTTAAATAGAAATATATTTATACATCTGATTCTTCAACATCTAAGCACTCGTATAGAGTGCTTTGAATATATTCTAAATACATTGGCATGATCCAAGTTTTAAATCAGCCAA
AAACTTGCTCCAAAAACAGGAAGATTCTACACAAATCGCCTCGCCCTTAAAGTTTCTGCCAAGTATTTGGGTGCGTGCGGATCAGTTC
(SEQ ID NO: 37)

Exon: 2..134
Exon: 208..388
Exon: 2321..3216
Exon: 3631..3664
Exon: 4521..5472
Start ATG: 2

Transcript No. : CT18539
ATGCGCAACTCGACAAACATCTTCCTAACCAACCTGAGCATTGCGGACCTGCTGGTGCTGCTCGTCTGCACGCCCACCGTCCTCGTGGAGGTCAACACTCGCCCGGAGACCT
GGGTACTTGGCCACGAGATGTGCAAGGCTGTGCCGTTCGTGGAGCTCACGGTGGCCCACGCCAGTGTGCTGACCATCTTGGCCATCTCGTTCGAGCGCTATTACGCAATCTG
CGAGCCATTAAAGGCCGGCTACGTCTGCACCAAGGGGCGGCCATCCTCATCTGCGTCCTGGCCTGGGGCATTGCTGCGCTCTTTACGAGCCCCATTCTGTGGGTGGCCGAG
TACAAGCTGGCCGAGTACATTGATGGATCGTCGGTGGCCGTGTGCCTGACCCAGGCCATCAGCGACTGGACGCTGGCCTTCTTCCTGATGACCATCTCGGTGTTCTTCGTGG
TGCCGTTCGTGACCCTGGTGGTGCTGTACGGCATCATCGCCCGGAATCTGGTCTCCAACAGGGCGGCCATGCTGCGCGCCCGTCCCACGAAGCCGGAGCTTAGTCTGAAGGC
CCGCAAGCAGGTGGTCCTAATGCTCGGCGCCGTTGTGCTGTCCTTCTTCGTCTGCCTGCTGCCCTTCCGCGTCCTCACCCTGTGGATCATCCTCAGCACGGACCAAACACTG
CACGATCTGGGATTGGTGCGCTACTACAGCTTGCTGTACTTCTGTAGGATCATGTTGTACCTCAACTCGGCCATGAATCCGATTCTTTACAACCTGATGTCGACCAAGTTTC
GAAGGGGTTTCAAGCGGCTTTGTCAGGACGCAGGGCGATTACTGCTGGAATTGGTGACATTGGGAAGAAGGAAGGAAGACTCTTCTCGCGGGCGCAGAGGTACCTTGTCACT
GGGCATGGGCACCAATACGAACACGAATACCAACTCCTCAAATGCCACAGGAGCCACGAGCTCCAGCATTCTCTCGAGAAGCTCCAATCGCAGGTGCAGCGAGGATATCAGT
CGCACCCGCCTTAAGATCGAGATGCAGATGCCATGTGGCAGCGACCTGGAGGCCATGGCCATGCTGCAGCATTCCACCTTGGGCAAGGGAATTGCCAGGCGAGTGAGTGATA
GCCGTCTCATGCCTCTGAGAAATCATCAACCACGACGCCACAAGCCGCAAATAAGTTTCGACGAGGAGTCACTGGAGGAGAATAAGCGTAGGTTAGGAGTCCAGATATCGAC
TGGCAGTAAGGGTCCCATCATCGCCATATCCACCTACAGCGTGGAACCCTACGGGGACGGAACCGATGCTCCCGTTTGCACCACCGCCGCCGATGGTTTCTGGTCGATCTTC
TACTTCGTGGGCTGCATCACGGTGTTTTTCTTCCTGCCCTTCGGCATCCTGGTTCTTCTATACGCGGCCATCGCTTACAAGCTGCTCCGTCCCAACAACGCCTTCCACCGAC
CAACCTCCCCGCAGCCACAGCAGCCGTCCGGCGGAGCCACCAGTGGCTCCTCACAGGTGCCCAGCACCAAGGGTAACAGCCATCAACAAAGCAACGGGATGAGGAAGCATCG
CAAGCAGGTAATCTTCATGCTGGTGGCCGTGGTGTCTAGCTTTTTTGTTTGCCTTCTACCCTTCCGGGCCTTCACCCTCTGGGTGATCCTGGCCAGCGCCGAGGATGTCGAG
GGTCTGGGCATTGCCGGCTACTACAACCTGCTGTACTTCTCGCGCTTCATGCTCTACCTAAACTCAGCCATGAACCCGATCCTCTACAACCTGATGTCCTCCAAATTTCGCA
GCGGCTTCTGGCGGCTGCTGCTCACTTGTCTGGGCCAACGGCCACATCACCACCATCGCCACCACTACCATCAGAGGCAGCATCCAACGGCAGGCGGAAGTGGGCGCAATGC
GTCCACGCGACAGGAACAGGATGCCGAGGAAGGAGCTGCGCTGGCGGGAACGACCAGCGCCCGACATCCACGTCGCACACTCCGCCGCGAGGCCACCTTCTTGATCAACTCC
ATATCCACCTCCTCGGGTACGGATCGCACCACATCATCATCGGCGTGGCGCAGCAACAGTCTGTCCATTTCCGGTCTGAGCGAACGGGAGCGCGGCATACTGGGAGCCGCTA
TCATCGGCACAACGGCTGCCACCGTTACAACCGCCTGTCTGCAGGAGCGACGCGCCAGCAAGATCTGA
(SEQ ID NO: 38)

Start ATG: 1

MRNSTNIFLTNLSIADLLVLLVCTPTVLVEVNTRPETWVLGHEMCKAVPFVELTVAHASVLTILAISFERYYAICEPLKAGYVCTKGRAILICVLAWGIAALFTSPILWVAE
YKLAEYIDGSSVAVCLTQAISDWTLAFFLMTISVFFVVPFVTLVVLYGIIARNLVSNRAAMLRARPTKPELSLKARKQVVLMLGAVVLSFFVCLLPFRVLTLWIILSTDQTL
HDLGLVRYYSLLYFCRIMLYLNSAMNPILYNLMSTKFRRGFKRLCQDAGRLLLELVTLGRRKEDSSRGRRGTLSLGMGTNTNTNTNSSNATGATSSSILSRSSNRRCSEDIS
RTRLKIEMQMPCGSDLEAMAMLQHSTLGKGIARRVSDSRLMPLRNHQPRRHKPQISFDEESLEENKRRLGVQISTGSKGPIIAISTYSVEPYGDGTDAPVCTTAADGFWSIF
YFVGCITVFFFLPFGILVLLYAAIAYKLLRPNNAFHRPTSPQPQQPSGGATSGSSQVPSTKGNSHQQSNGMRKHRKQVIFMLVAVVSSFFVCLLPFRAFTLWVILASAEDVE
GLGIAGYYNLLYFSRFMLYLNSAMNPILYNLMSSKFRSGFWRLLLTCLGQRPHHHHRHHYHQRQHPTAGGSGRNASTRQEQDAEEGAALAGTTSARHPRRTLRREATFLINS
ISTSSGTDRTTSSSAWRSNSLSISGLSERERGILGAAIIGTTAATVTTACLQERRASKI*
(SEQ ID NO: 39)

Name: Thyrotropin-releasing hormone receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000012807287
CAGAAAGTGTAGGGAATGTAGGGAAATGCAAAGCGAGACCCTTTGGCTGCTCACTGACTGAAATATCACCAGACTCGGCTCGGTTCTGCCGCTCAAGTGGCAGTAAATAAAA
GTTGTGAGGAGCCTTGGCGGCATGAGTAATAATGACCGAAATCAGAGCACCATCACCATCACCAACCATCGCTACCACCGCCCAACGCGTTGCTCATTTTATTAAAATAACA
AAAATATTTTGCCGCTCGCAGCCTTTTTACGATGCCGAACCGACTGGCGGCCCATTTGTTATTTGGCGGTCATTGCCACCAAATGGTGTCCTGTCGCTCTTTCTCGTCCTCG
TCCTCGTCCTTGTCCGTTCCACTGGGCCAACTCGTGTGTCGTGTGTGTTTTGACCAGAGTTATGCAGCTGCGGATGGGCAAAAATGTTTTATTGCCAGCTGGGGAATGCCAA
AGTGGCAAGAGGCAGATTGCCCGATCTTAGCTGTTTAAATAAATTTAAGAGAGAGAGAGAGAGAGAGAGAGATCTGGCACGAAGCACTCCATCACTGGTCCAGCGATCCGTTAA
GCCATCGAGGGTCCTTCGAATGAAATGCCAGCGCTGACAAGGCAAACACTGGAGATTTGTATAATAAGGGGTGGGTGGTTGTGCCGAAATTTGGCACTGATAGGCAGGACAC
TTGTCAGGCTTGGCGCAAACAAAACTGACATTTATATTGATGTGTCCCTGTACCCAATTAGATGCCCATCATTATGAGAGCAGGCTATCTCTGCTCTCGTGTCCTTCCCCCT
TCTATTTTGCGCCCGGGAAGTCATATATCGAACATTAGGCTCTGTTTGTGTGCAGAGCAATTTGGCGGGAAGTGGCAGCAGGAGGACGCACATGATATTAGTAATAATGTAA
TTATATGGTCATAAACAAAATAATAAATCAATAAACAACACACACGCACACGCAGGCGGCACACGAGTAAAAGTTATCACGGGGGGAAAAAAATGAAGGAAAAATGGGAGG
CAGAGCAGAGGAAGTGAAGAAAAGGGAACCGGAAAGGAAAGCCATTCATTTGGCCCCGCGACTCAAGTGCCGTTGTCTACGTTTGCGAATGTGTGTGGTTGTGTGCGTGGCA
GGACTTCCTGTTGCCACATGGACGGAAACGGCGCCATAATTATGCCATAAACCCAACCATAAAGGACATTTTTATAGATGTGCCCGAGTTAATCCCTTTGTTGTTCCACTCG
CTACCGCACCCATCTCTCATGTCCTTTTTATCCGCACATCCTTTCAGCTTACTCCACCGCCGCCCTGCTGGCCATCGTGTTCGCCATTCCCTTCGGTATCCGCATGCTGGTC
CACAAGGATCGAGGACAGTGGGAGGAGTTCGGCCCAGCCTTCTACACCGCCCACCTGGAGCTTTATCTGGGCAACGGCTGTTTGGGTGAGTCTAGGATTTTGGATTTGCAAC
AGGAGGAGCTTTCAAGGGTCGAAGCTCTTCTAAAAGGTTCAAAAGTTGGACTCAGCTTTAAGTTCACTCGGATATACTAAAAATATATATAAAACAACACGTTATGAAACTG
TTTATGTACCATAAATGGAGCGTAATGTGCTTGATTTAAAGGAACAGATGTCAACAGTGCTTAAATTGTAATAAAAAGCTATTTGTTTTAGGAGCTAGCCAGACATTATTAG
TATGGACTGATAAGCATATGTTAGAAGATAATATTTCATACATTTAGATACGCCTATAATTTCATAATAATAAATCATAAAATTTATTCCGTTGGGATCATTTAATCCTTGA
ACAGATATCTGTTATTTGATAATATTTACAAGGAATATTATCCCATTTGCACTCCTAGGCGTTGGCGTAATGATGTTACTAGTGCTGACCATTGAACGTTACGTGTCCGTTT
GCCATCCAGGATTTGCTCGACCAGTGATGGGACCACCTGGGTAAGTCGCATTGAAAGTCTGTAGATGGTCTATTTAATTGATCCAACCTAGCGTTGTGGTATTCCTCACCTG
CTTGGCCACGGTAATCGTCTATCTACCGAGCATCTTTCGTGGCGAACTGATCAAATGCATCCTTGGATCGAGTGACGTATATGTATATTTGCGACGCGACAACACTATCTAC
CAGCAGACCATCTTCTATCGCGTCTACAAGATCATGCTGGAGGTGATCTTTAAACTGGTGCCCACTTTGGTGATCGGTGGCCTCAACATGCGCATCATGATGGTCTACAGGC
GGACCTGCGAGCGCCGCCGCAAGATGGTGCTCAGTCGTCCGCATGCCCAGGGACATGGTCACGGACATGGGCACGGACATGGTCATGGACATGGACATGCACATGGACACGG
CTATTTGAAGGACGACGATCCGCGAAAGTTCGCCGAGGAGCGGCGCTTGTTCCTGCTGCTCGGCAGCACATCGATTTTGTTCCTGGTCTGCGTCTCGCCGATGGCCATTCTC

CACATGACGATCGCCTCGGAGGTCTATCCGAGCTTCCCCTTCCAGGTGTTCCGTGCCAGCGCCAATCTGCTGGAGTTGATTAACTACTCGCTGACCTTCTACATCTACTGCC
TGTTCAGCGAGGATTTCCGCAACACCCTGGTGAGGACGATCAAGTGGCCCTGGTTGAAGGGCAAGTTCTGCCACCAGGCCGAGCACGAGGTGAGTGCCAGTCCACCGGCCAC
AGCCGGAACCGTAGCGGTAGCCGGAACCGGAAATGGCCACGTTTCTATTTTTCATCCGGCAATACCGGCGCTGACTCTCACACCCGCAGAACCCGACGAACGGCCCCGGTGT
GCCAATGGCGTGCTTCACTAAGGTGGACAGGGGCCATCAAAAGCACCACATAACTACCACCTCGGGCCTGGGACGATCGAGTAGCATTTAGACAATGCCTACCACCCAACAA
CAACACCACCACAACCCATCCGCTACTCGAGATCCGAACTCCGAACACCGAGAACCCTCCCAGCCCAGAGTAACCCCACCACCACCACCCAACTCCAAAACCGTCATCACCA
CTTCCACTTACCCCCCAGGTCGTCGAGGTTAAAGGGTGCCCTTTCTCTCGCTTGAGATCCTGCGACATTTGAACGTGTTCAAGTATCGATGAACGAAAAATGATATACGATA
AACAACTAACGATTAACGATAATTGATATTCAGCGCCACGCATTGCGTATATAAGTGTCTGTGTGTGTTCCTTAAACATATATATTTAAATATATGGATAAACATTTTTTT
TTTTTAGCACTAGAATTAGGTACTTAAAAACCAAACTTAATCGAAAACACATCCGTAATTGGTGAAATAATTGATATCGATAGGTAAAACCAGTTCGACTTTCGATCAGCAT
TTGCGCCGAATGATTTCACCACAATGTAGCTTGCGGAATGTGTATGTCGCAGGATCAAGATCCTCAAGCCACACGTATAGATGCAGATATGATAGTTTTATCAATGGATAGA
TTGCTTGCCGTATTGATAACCAAAGAGCTCGTTTACTCTAGCATTAGATGCAGACAGAAACAGGAACAGCGACAGGACAACGACAGACAGACATATTGACATATAGACAGGT
ATAGACAGGTGAGATGGAACAATAGCGAGAGAGAGAGAGAAAGAGTTTTACGATGATAATATGCAGACATACAATACTTATATATACACTTGCATATGCCTATATGTGAATG
ATATTTATGTAAACGACGTATTTATGTATACACGCGTAACTTTAAGCGAATATATTTATATATGCATATATGTATCTATTGATTATAGCATTTAAGTGCGCGCAAGAGGAAT
CCGGATTCCAACC
(SEQ ID NO: 40)

Exon: 1001..1088
Exon: 1280..1429
Exon: 1851..1944
Exon: 1996..2821
Start ATG: 1001

Transcript No. : CT18637
ATGGGAGGCAGAGCAGAGGAAGTGAAGAAAAGGGAACCGGAAAGGAAAGCCATTCATTTGGCCCCGCGACTCAAGTGCCGTTGTCTACCTTACTCCACCGCCGCCCTGCTGG
CCATCGTGTTCGCCATTCCCTTCGGTATCCGCATGCTGGTCCACAAGGATCGAGGACAGTGGGAGGAGTTCGGCCCAGCCTTCTACACCGCCCACCTGGAGCTTTATCTGGG
CAACGGCTGTTTGGGCGTTGGCGTAATGATGTTACTAGTGCTGACCATTGAACGTTACGTGTCCGTTTGCCATCCAGGATTTGCTCGACCAGTGATGGGACCACCTGGCGTT
GTGGTATTCCTCACCTGCTTGGCCACGGTAATCGTCTATCTACCGAGCATCTTTCGTGGCGAACTGATCAAATGCATCCTTGGATCGAGTGACGTATATGTATATTTGCGAC
GCGACAACACTATCTACCAGCAGACCATCTTCTATCGCGTCTACAAGATCATGCTGGAGGTGATCTTTAAACTGGTGCCCACTTTGGTGATCGGTGGCCTCAACATGCGCAT
CATGATGGTCTACAGGCGGACCTGCGAGCGCCGCCGCAAGATGGTGCTCAGTCGTCCGCATGCCCAGGGACATGGTCACGGACATGGGCACGGACATGGTCATGGACATGGA
CATGCACATGGACACGGCTATTTGAAGGACGACGATCCGCGAAAGTTCGCCGAGGAGCGGCGCTTGTTCCTGCTGCTCGGCAGCACATCGATTTTGTTCCTGGTCTGCGTCT
CGCCGATGGCCATTCTCCACATGACGATCGCCTCGGAGGTCTATCCGAGCTTCCCCTTCCAGGTGTTCCGTGCCAGCGCCAATCTGCTGGAGTTGATTAACTACTCGCTGAC
CTTCTACATCTACTGCCTGTTCAGCGAGGATTTCCGCAACACCCTGGTGAGGACGATCAAGTGGCCCTGGTTGAAGGGCAAGTTCTGCCACCAGGCCGAGCACGAGGTGAGT
GCCAGTCCACCGGCCACAGCCGGAACCGTAGCGGTAGCCGGAACCGGAAATGGCCACGTTTCTATTTTTCATCCGGCAATACCGGCGCTGACTCTCACACCCGCAGAACCCG
ACGAACGGCCCCGGTGTGCCAATGGCGTGCTTCACTAA
(SEQ ID NO: 41)

Start ATG: 1

MGGRAEEVKKREPERKAIHLAPRLKCRCLPYSTAALLAIVFAIPFGIRMLVHKDRGQWEEFGPAFYTAHLELYLGNGCLGVGVMMLLVLTIERYVSVCHPGFARPVMGPPGV
VVFLTCLATVIVYLPSIFRGELIKCILGSSDVYVYLRRDNTIYQQTIFYRVYKIMLEVIFKLVPTLVIGGLNMRIMMVYRRTCERRRKMVLSRPHAQGHGHGHGHGHGHGHG
HAHGHGYLKDDDPRKFAEERRLFLLLGSTSILFLVCVSPMAILHMTIASEVYPSFPFQVFRASANLLELINYSLTFYIYCLFSEDFRNTLVRTIKWPWLKGKFCHQAEHEVS
ASPPATAGTVAVAGTGNGHVSIFHPAIPALTLTPAEPDERPRCANGVLH*
(SEQ ID NO: 42)

Name: G-protein coupled receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384543
AATAATCCACCTTCGAAGGTAAACCATCTTAATTATGCTTTAAAATTGCTGGATTATATAAACCCCTAAGGTAGTTCATCCCACTTGTTCAACTGAAATCGACCCACACAAA
TGGTGGTTGGTATCTATCTACCAGAGTTAGGAGCTCGTTTAACCCAATATTGCCTATTCGTAGAGATCCGCATTCCTGCGAGGAGAAGTCGAACTGCCGTAGCTGAGCTGCG
AGTCGGTCTTCTTCAGAATGGAAACGTGGTGGGTGGGTGCATCACCCGTCCGGTTGGCCATCAGCGAGTCCTGTTTGTGGAGCAGGCCGCGTCGCTGGTCGGCGAAGGAGAT
GACCGAGTCCTGCTTCACCAGTTGATGGCGCCTTTGCTCTGGTCGCTGGTCCACAAACGAGATCACCGAGTCCTGCTTTACGAACTTCTGCCGCTTATCGAGCAGAGAGCTC
GGTGTGGCGCCCTGTCGCCGGCCCACATCCATGGAAACGCTATCCCTGTCCACCAGTGTGTCGTTGAGACCGTGTCCTGCCGACCTGCGTGCCTCCATGTAGGACACAATTG
AGTCCTGCTTGGACAGGGAGACCTTGGAGCCATTGCTGTCCTGCTGGCAGAGGGGTCGCCGGATCCCAGAAGCATCACTGCAGGAGCTCTCCTGCTTGATGAGGAACTGACG
CGATCCGTGAGAGTCCTGCTTAAGCAGCATCTGCTGCGACAGGTTGCTGTCCTGCTTGAGCATCATGTTGTGGCGCCGCAGCGAGTTGTGCTCGTAGATGTAGGACACCGCG
TTGCCCTTGCTGCTCGGACTCTTCCTCGAGCTGTTCTGGTGGTAGTGCGATGACTTCCGGGTGCTGGTGGTGCTGGAGGCGATGGCCACGTTGTCGTAGTACCTAGGTGGTT
AAGGTGGAAAAAGGGTTTTATCGAAGGGAACACGATTTCTTACAAATCTTCATATCAATTTTGAAGGCCTTCACATATTTTATCAATATATTTGAGGTTAGAAATTAACTA
ACCTGAGGCGTCCGCAGCAGCTGCGCGACTTGATGATGGCCGCGAATCCGCGTCGGTACTTCTTGTTCACTGAGTAGAGTATGGGGTTGATGCAGCTGTTCGAGGAGCCCAG
CCACTGGGCCACCGGCATCACCTTCTTGAGGATCTCGAACTCCTCCTGCGAGATGTCCGAGCCGAATTTGATCCGCGCAAAGATGACATAGAGCGGCAGCCAGGACAAAACA
AACAGGATGACCACGGCCACCAACATCTTGATGACCTTCACCTTGCTCTTCTGCTGCATGCGGTCCATTTGCGCGTCCTTGGACTCGCCGGGGATGGACCTCGTGGAGACCT
TGATCCAGATGAGCACGTAGCAGAGCGTGATCAGGGACATGGGCAGCAGGTAGCAGGCCACCAGGTTGGCTAGCAGGAAGTAGAGGTTCCCATCCGTGCCCGGGGGCCACAC
CTCCTGGCAGAGGAACTGCGGCTGCGAGTAGGCCGAGACCAGGGCGTCGGAGAAGACCTCCTCGGCGGGCACCAGGTCGAAGAAGAGCAGCCAAGGAATGGTGGTCACCAGC
GCGATTACCCATATGCCGATTATCATGATGCGGGCACGTCGCTTTGTCATCTGCTTCAGGGGCCACCAGATGGCAATGAACCTGCAAGAGGGGAGTCCCAACAAAGGGGATT
TTTACCAAGTAGTCAAAGAAGGAGTCTTCGTAGGAGCAGCAGCAGCAGCAGGAGCTCCGGCTTCAGATGCCACTGACGGCTTTGAGCCGTCAAGGATAATGGCAATCATCAA
GCTGTCGAGTCGAGTCGAGGCGACGATGACGATGACGACACCGTTTCAACTAGGAGGAGCAGCTCGGCATGCAAATATCTCCGCTGATTAATTCGCACACGCAGGATGCCGA
TGCCAGCAGCCATCCTTAAGCCTTTTTCTGCTGAGTGTTCGCTTTCCAATCTCGACTATTGAACCCGCCTGCTACTCCTGCTCCTGCTCCTCCGCCCCAAAAAAGCGGCGCT
GCGGATCCTGTCATTTTTTATTCTGCTTGCCATAATGAATTGTAATTTGAATAAGAGCTGCGAGGTAGTGAGACGAGTCGGTATGGGGCCAACAACGAGTTCAGCATTTTTC
ATTTAGCTTAGTGCAGCCGGCTCATTGTGGCGCCCGCTCAGTGGTCATAGACAAGACTTGTCCCGGACGAACGAAACTGGTTGTCTGCTCCTGGCCCGAAACCATGTAATTA
CGTGCATTGTGTTCGTATTTCGACTTGGACCTTGGGTTGTCGCCTGCGTCTGTCCGCCAAAACATTTGGCAGATAACACAGCGGTAATTCCATTAGAGGGCGGCCATGCCA
GGCCACTACCCCAGAATCCTATTCCTCACCGACTAAAAATAGCCAGCTTAACAATGTGGCCGAGACCCTCATCTCTGCGAAAAAAAAAAGGTGGGCCGCCTGGGCGACCAAA
CAAAGCTGTCCGCAAAGCCAATAATCAAATTGTCATAACGTTTATTTATAGGAGCAAACCAATTGCATTTGCATAATTTGCCCATGCGACTTCTATGGTCATAAATCAGGCG
AGAGCTTAATTGAGTTCTTTTTGGGTAAAAACAGCGGAGTCGGAAAAAGTAATCAGCAGTTTTCGCTGGTGCAGACTAATTCAAACTAAACCTCTAATTTGATGTTGAAATT
TTGTTAAATCGTATTAACTACACTTTTGATGATGTAAACCATTTTATGAGCTTTAAAAACTAAAATTTCTCTTATGAGTTCTCTTATACGATTTATGTATGATTTGTTTCAC
AAATAAACTCTCCTTATAGATTCAATTGCTGTCGGAAAATTAAATTATTAATATTATTTTGGTAATCACGGGTAATTGTTGCCCATCAATTCCGTTGAGTACCATATGCGTA
AATCTCTAAGGACGTCAGTTCAGTATTGTTAGCATATTTATTAAGAAAAATCTTGCCACAATCAACTTAAACAAAGTAGTTAATACTCGCGCTTTTAAAGCATTTAATAAGC
CGCAAGTGTATTAATTTAATATTTGCCACACAATCCTGTGTTTCTGTCTAGATCCACGTGGTTTTATGGCTCCTATTTCGCTTCCTACTGCGGTTACCCATACGCACTGTTG

ACATTTGCGACGCAAAGAAGACACAAAAGCCGCAATGCCAGCACCGTAAAGACACAACTATGTCTATAAGCGAGTGTGCTTTCTTATAGACTTTATGATTATAAGCAATCTA
GTTATATACATATACGTATTCAGACAAGAAAAGACCCTAATTAAATGACAGTGGCATAAGTATAAATATAAGTAGAAGTAGAAGTATAAGCAAGTGATAAGCTAAGTCGAAG
GGGCATTACGGGCTCTAAGACCGTGGAGAGATTTAAGTTTCATGTTACAGATAATAAATCTTAAATTTCGATTCAGTTTACGATGGCTCGTAAAAGGAGAGAATAACAGCAA
AACAACTGGATAAACGGACATCAATGACGCTGGACTCGGACTTCGACTGGACTTGGAAGTGGAAGTGGAATTGGAAGTGGATGAGGATGTGGATGTGGATTTGGACTCGGT
TGGGTTGGCGTTAAGCGTGGGCGGAGGCTGACATGCGGGCCTGTGGGTGACCAAACGATCAAGTGAGGCGCTATGCCTTTGTCCCGCTGGCTGAAAAATGACAGTGGCAAAA
ACAGAAGCTTCTGGCCAAAACTCCTCCTCTGATTGTGGCCAGAAATTGATATAAATGAAAGTGCACATAAAGCGACGGGCGTCTTATTGTCCTTGCGGCTATGGCTAGCACA
ATTTGAGGGACTTCAGCATCACTTGATAAATCTGTTGACCGCGACATTTGCCGAATTTATTGCTTTTTTATACTTTATGGGCCAGGAAACGCAGGGCTGACACATCGATTCC
TGGCGGCCAATCATCAATGGTATCAATATTCGCGGTATCGCGGCCAACAACATTCTAAGGCAATTGAAATGCACATAAAATGAGAAACATGAAATTGATTGTAAGCAGCGCC
TAATACGGACAAGGTGAGCCATAGAAGCATATGATGATCTTGATCTTGAGACGACTGCTCCACTGACTCTCCTACATAACTGCCTAACTGGCTAACTGCCTTACTGACCAGT
TTTCCCCGCAGCTGATTGCATTGAGTGGCGCTCACTTTCGGTTCGGTCGCCTGGTGCAAATCCACTATAAAAGCCGCACTCGGAGCTCAAGTTGTCACCAGTTGACAATTTG
AAATCATTAGCCTAGCTACCTTTCCTACTTCAGTTATCCCGCCAAAATCCAACGCCTCAAAATGTTCAAAGTAGTGAGTATCATTGTGCAGAGAAGAAGTTGTAACCAGTTTGA
ATCCAAGCATCGTATCATTATGATATTATATCAGTTTGCTCTATAGCATTTAAAAATAAATGCAATAAATTAAGTTTATATTACAACTCATATGAGTCACCTAGTTTGTTGT
TAGGATTGGTCTTATCTTACTCAGAATGAGTAATATATGCTTTAAAGTCTGCGACAACTAACCAAATTCACTGTGGAGTGGCAGCCACCTGCTGCTCCGTCTAATGCCACTA
ATCAGGCTAATTCGTGAACTCTGATCCGCCGCAGCTGTTCGTGCTCGCCGCCTTCGTCGCCTCTCAGGCTATCGCCCATCCCGGCGTGGTGGCCGTGGCACCCGTGGTGGCG
CACCCGGCGGTGGTCCACACGCCCATCATCCATCATGGAGCCCACTCGGTGCACTCGTAAGGCCATTCATGTCCAATCTGCTTATCACCACCGAACAACTCACCCCTTTCAT
TTTCAGCCATGTTGTTCATCATCCGGCAGCCGTCAAGGTCATCACACCCGTCGTCCACAAGCCCGTGGTGGCGGTGCATGCTGTCAGGCCGGTGGTTCCATTGGTTCCAGTG
CATCATGCCGCCCCCGCCGTCGTCGTGCATCATTAAGTGCAGCGAATACGGTTCACTTTCAATACCTTTCCCGCTCCCGATCCCGTTCCCGATCCCCACCAGACCCAGACCA
ATCCCTTCCTGAGCACCATAACCACGATCCCACAATAAAAGTTCATAGCCACGTAATTGGAGCGAGTGTTTTCAACTTACCTGTCCAGGGAAACGGCAATCAGGCTGTAGAC
CGAGGCTGCCACGGAGACACCTTGTATGTAGGGCACAAACTTGCACATCAGCCATCCGAGCATCCAGGCTGCGAGTTCCAGCAGTTCGCGAAACATGAAAGGGGGAAAAGTG
CGGGAGGTGAGTAAATATTGGTCATAAACAACGCACAAATTATGCTGAGTAACTTTCGGTCACATTGTCCTTGCTGGTGTCCTAGCCATACACTGAAGGAAATGTGAGGGCC
TATGGCAACAAAAAAAGAAGAAGATGCAGATTTTTCTAAAATACCCCAAGGACACTTTGCTCGAAAACTTTATAAATATCCTAACCACGAAAATCAAAGCAAATTAATGTGG
TTAACCGACCGAAGGATGCTTATTGTTCTCATGGGAAATATTCGAAATAGTAGATGGTGTATCACTTTAAATACAAACCGTAAGTCATCTTGACTGCTTAGCTAAGAAAAGA
GTGAAAATATTATAAAACTAATAATTCTCCGCTAAGTATTATAATATATTTCATATCTTACTTTTATTCTTGAATAAGTACAGTTTTACAGCTGGGGATCATTGAGTTCTTT
ATAGCTTTACTGCTTTTCCCCGAGTGCACTGGTAGTGGTTATGCTCATGGTTGTGGTGATGGTGGTGGTGGTGATAGTGGTAGTCATGGACAGTTGGGTTCGAGATCTGGTT
AGCGTTCAATGCTTCTTGTCATTGCCGCCGCCCACGTTCCCGTCGTTGTTTGCTGACTGAGATTTATGTGCTGCATATCATTCATAATTTGGTGACAAGTCTCAGATTGCGC
CCAGCCATCGTAGTGGGTGGGGCAGAATGTGGAGTCCTGGGGGGCAGAGCATTGTTTGCCCAGACCCAGACTCAAAGCCGCGGTGGACCGCAGGACTGCCTAACTTAATTTG
TATTCCTTTTACTCCATTTTGTGGCTGGATGATGATGGTAAATGGAATACACGGGCCCTACTTGCGCATAGTTGAGCGTTATGAATTTTATGCTAAGAGGCATAACTCACCC
AGACACAAAAGGTTATAAAGCATACATACATGGGCGTGCAATGGGAGGGGCAACTTCCCAGGGTTGATGACAGTCGGGCGGAGAAAGGCGCCCGGTTCATGACATATCGAAT
GGCAATTAAACAAGTGCAAGTGCGGCCCACACATCATTAGCAGTTAGCTGTTTGGCAGTTTGGCAATTTGCAGAAGGAGTTTACCGAGTTCAATGCTCGACCAGGAGACGCC
TCCTGCGTCATGGGGCTCCTAATTTCCTGTCTCGTTCGCCTGTCTTATCTCTATTTGTTGTCGCTTTTTATTATGCAGACGCACCGTAAGCCGCCAAGGCAGCCCAATAAAA
GCTAACTTTAATATGATTTGGCACATAAAAATATAAATATTATTGGCGCTGTGGCGCTTCGTGCAAGTGAACTCTGTGGCTGTGGACAGGTTTTGGTTCTGGACAGGATCAT
GTGTGTCTGCTGGGAGGACATACACATCCTCCCGCCTGGCAAACAATCAACAGTGTGTACACAGGACAGCAGACCGAATTGGTTTTTCGGTTTTATTCCTCCGATTTGAAAT
GGCTTTCGGGCAGACCTGCAAGTGCAACAGACTGGATCCGAGGCTATTAATGCGAGTGTATTTGGCCACTCACTTGATATTATGCTGATAAGGGTGAACTTTGATATATAGC
TGTCTGCGGGAGCTCTTGACTCGGCAAACTTTATGCTTTCAAGTTGTTTGAAAACAATTTAAATTCAAATTCGGTTCCGAAATGCTTCATCACTTGCGGGACACCGGATTAG
CAGTGCCACTGGGTGAACTTTTTTTTTTTTAAATACATGTTCTTCCAACTTCCCTTCGGCAAATAAACTATCTAATTGTGTATGCAATTCTATTAGGACAGGCAAACGTG
ACTCGGTTGGTTTGTCTGGGATTAAAGTTTGCATCAGTGCACTTTGGCAGTCAAGTGTCTTAACCAAGTTTACATTGTGGTTAGCTTAGCGCAGTCAAGTGACCACTTGATA
ATCCCGAACAAACAATGGTTTTCAGTTCATATGGAAAGTATTTTGACAGGCTGAGGCATTCAGCACGGTTCAACTTATTAAATTGTCCAACTCGCTGATTTGGATATTTAAT
ATAAAAATGGAAAAACCACATTTTACAGAACCAACGAGAAAAGCTGACATTTAAAATTGGTTTTTGTTAACATAAAACGATTGAGAAGGAATAATACATAATTAATGAACCG
CAAAAACCGCATTATATATAAAAAAAATCAAATAGGTCAAGGGTTATGCCAGTAGTTATTTATAGATACCACAAGCTTAGGTAAGACATGATTATTTCAAGGAACACCGAGT
CATTCGAAACCAATCTTTAATTAACCCCAACCTTAAGGCCACATTAATGTGTTGTACAAACATAATTGTAGCTACCGAAATAACCTAATTGCCTTGCGTGACGGGGCTAACC
CGAACCTCCTTGTGGCATTACCCATTTGAGGACTCAAGTGAAACCAAACACTCCACAGTTTGCTCGGCCAACATCCAGCCTCCCCGTCAAGGTGAGAATGGGATGGGTGTTT
GATCCGACTTGGAGTGCGTTTGACATGGCAACATATGGCAAACAGACGCCTCCATTTATACGCGAAAGCTGGTGCTCAATATTGATCAACTAGAGGAGATAATACCAATCGA
TACTCAGGCTTATCGCCAGCCTTCGATTGTGGGTGGTGGGTGGTGGGCTTGTGCGCCCCATCAGCCCTTAAGAAAATGTCAAGTGCGCCCTAAGCCTTTTATTTTTTAGCCT
CTTACCCTCCTTACAACGAGAGCCTTTGTTTTTCGGTTTCGGTATATATGACTTTTAGAAATGTGATTCAACTGAAAATATTTACCTTATTTGATAATTTGCCTGCAGAATT
AAATTTAACGCGGTTATTTATGAATAATGTCCCCTTATTTGGAAACAAATTTTTCTCAAGGACAACTTTAAAACGCGTTATCTTTTACAGCAACAATAAACATTTTCTCTAC
TGTGCTTTATGAGGCTTTAATAGCGCCAACAAAGTTTTGTTACATTTTTATCCCCAATACCGTGGGTGATTGTTCCTATGTAGGCAAAGTGGATTGGATTATTCCGGGCATT
GCGGATTCGTTTCGGTGGCGGCATTGTCATGCACGGAGTTTGATCGTAAAGTTAATGAAATAATCAGGCGTGACAAATCCGCAGGTTTGGACTTGGGAGTGGGAACGGGCGT
GGATTGCAGTGGAGTGGAGTGGAGTGGAGTGGCCATGGAAAAGGCATAACATAACACTTCTGGCAAATGCGTTATTGTTATGATTATCGCCGCGGCGTCACGTCCTGCGGCG
ACTTCTGTGGCAGCTCTTCGGGAAAATTGAATACCACAGCTAGTTCGCAGAGAAATGCGCCTGGGACGCGGAGTGGGAATAGGAATTGGAATCTGGGAATCGGTGGTCCTCG
ATGTGAGATACGGGATTGCGGGAGTAAAGAAGCTGCAGATCTTAGGTGGAGGGCTCAGAGGTTGGATGCCATTTGCCGACGGCACGTAGAGACTAAATTAAAATACAATTGG
CGAGCTGAAAAATGACTTCATCAGGGATAGGGAAACAGAAATGCGGGAGAGTCTGCCTCCAAGTTACGCATACGCCGCGTTCGCCCGGCACAAATCACGTGGCCCCCTAATG
AACAAACTGGCAACGTGACTGAAAGATTTAGCCAACGGGAGATGCTGATCTAAAAAGGGCCGTAATGATAACGATTCCCCATTGATAAAATAGCACAACAATGGCCATAAAG
CGCCATCTGAGACAATTGCCGTTGCTGTTGCAAGGATATCCTTCGTCTCAAGGCGGCATCGTTGCTGGGAAATCAATTCATTTCGCATTTTATTGGATTTCAAAGCACACAC
ATAAATAAGGGCCAGCGCCACATTCCGACCAACAACCACCACCCACCACCCACCACCCACAGGCCACCAGCCAGGAACCACCCACAAAGAGTACACGGCCCAACCCAATACA
GGTGCTCCCACACCCCACGCATACACACACACACACACCCTCAACCACACACACTGGCACACTCGCAAGTGTGTGTGTGCATAACCCTGACTATTAGGTTTGCCAAGCGAAA
TATTTAGGTCAATTCATCCAGCCGCAAAGAGCCTTGAATGCAAACAACCTTCCATTGCAATTCTGAGGCTTGAAGGTGTGCTTAGATAATGGCGCTGCCAGCGGCGATGGCC
GTGGGCGTCGCTGCCAGCTGCGACTGCGGCGGCGACGTCGTCTTTGACGGCTGGAATTTGATTTCATATAATTCATTTTGGTGGCAAAACAACCCGCCCACTTTGCCACTCG
CTGGGGGCGAGTAGTGGGAACTATCAGACAGCGATGTGAAATGCTCGATGTACAAATATGCTAAACTTCCGAACTCCCACGCACTGTGCTACACAGGGAGAAAAGGTTTTGT
TAGAGGAGAGCAAAGGATTAAGTGGTATGACTTTACACTGCTAATTTGATTATGAAGAATATTCGTAAGAGTTATAGCATTTCATCACCTGACGAAGTTCAAAAACATTTAA
GTTTAAAGCCAAAAAGTGTATATTTCTGAAAACGTGTGGATTGTTTAATGAAAATGACGTGTTAAACTTTTGTTTGTAAATAATACGCTACACATGTGTTTAGTTGGCAACA
CATGTTTTTTCCCCTGCATTAGTGCTATTTCACCGTTGCCGCTGTGAAACCGTTGCCTTCGTTTTAACTTCAGCAGGTTGTCGAAAGTGAATAGGAATGTAAGTGGGGTTGA
GAGTTGTGGGGATGAAGAGGTGAAGCTGAAATGAGCCACCCCAAAAGCCCCAATGCCAGCGGGAGGGCTGCAATTTGTTGTGACTGGCGTTTTTGCTAGGCATCGGCGGCTC
AGCTGAAGCTAAGACATGCCAAGCGAAAATTGCGGAGTGGGCGGCCAAAAGGGTGCCCCATCCAAACAACCATACAACCATCTAGCCCACCCAATCTACCATCCACCGCCAG
CCACCCACCACCCCCAGTGACCCAGTAAGAGCTGTCGGCCCTGCGCCCAGCTTGAATTTGTGAAATTAGGTCAGTTTTGTGTCGATGATCCTTTGGGCCACCCCGAAAAGCA
GCCACCAATGCACCTACACATGCACAGTAGCGAGCAAACACATGGGCGTGAGAGCGGGCGAGAGAGAGAGAGCGGAGAGAGTGCGCAGTGACACGGACACAAAAACGGAGAC
ACGGACACCGATACAGATACACAATGCAGGCGGAGGTAAACTCATTCACTTTTCGCCACTCGCTTCATTCGTTTCCAGCGATTCGCACTCGGCAGCCTTCGTTTACTACATT
TCCCCCACAGTGGGCTTCTCTGGGGGAGTGTATTGGTATGGGGTGGGATCGAGTGCAGTGCCGAGTCTAAGAGTAGAGCTGCATTGGAAAATAAATCTGATAAGTCAACTGC
ATCTGCATATTTGTTACCAGTTTGTAATATTAATATTCAGGTTAATAAATAATCTGAGGGGAAGAGGAACTAAAACTATTTTTATTATTATTATTATTAATTTAATTTAATT
TAATATTATTATTATTAATTATTTAAATAATTTTAATTATTTTCATATACTTGTGGCAAAGAAATGTAGTAACACAATTTGTTAAACTATTTCCTATTAATCGAACATATT
GTTTTACGAGTATGTTGACAACTCTTATTTTGCCAGCGTACTCGCTGTATAGCAACAAGAAAAGGAGAGTTCTCACAGGGCTGAGCCACTTAATTGGCAACACTTGGCACAA
CCTGCGAGCCTACGACCTGGCATTGGGATTTGGGACTAGGGACCTGGCACTGGGATTGGGGTAGCCATCGAAGGCAAATTGCAGTGAAGGATGGAGCTTGCATGGCGTGCTC
CACAATCGGAATTTGATTTACCAAGGATTTCGATCATTTTTCCATGTCAGCGGCGATTTGCATAGACATTTGCCAACCTGCTCACCTTGTCGTCCAGGTCTGCGACTATCAA
TTTGGATGTCAGTCCACAAAGTTCGTTTCGAATGCAAGCGGGGAAGTTGCGGAGGGGAAACCAACCAATTGGGGGTGTAAGACGGCAGACAGGGACAGATAGATAGATAGAT
AGATGGACGGATGGATAGGCATGTCAGGGAAATTTATTTAAATTAAAGCATACAATTATGCGTAGCCTCGTTTATCGGGTCGTTTGTGGGTCGTCTCTGGGATTTGGTGGAA
TTGGCCTCTAATTAAATGGACACTTGTTCTTTGGTCCATGCGGATTGTCGTTTCCAAGAAGGTAAGGGCTCCCATGGAGAGAGCACAAATTAATTTAATTGAGTGCGAACAT
CCGTTGCTACGGAAATCTGACACATTTCTGGTAAAGTACGCTGAATGAATTCTTCTTTTTTTCTTTGCTGCGTACTGTGTGAGATGTAGTTCCTTGATTTAATGAATAACTG

FIGURE SHEET 22

GCTTTTGTCTTGGCAGTCTATTTTAAAAGTTATCTTGACCATTGCAGAAAGCTATTAAATTTTGATACTAAAGCTTGGCATTTTCCAGAACAACTTTAATGATGCTGCATCG
TTATGTTCTTCTTGTGTATGTAGTAAAAATTCGGAAGAAGCCAGGGAGTCTGGTATATAAAATGCAGCTTACAAATAAACGGAAACACAGTATGTCTGTCAAAAGTAAATAT
AACCGAGTATGAATAAAATCCCATAATTTATTATTTTCCACCAGAAATATGAGCACTTTCGTTTGGACATTTATGCCGACATTATGTCTAATGAGGGTCAATTACACTTGCC
CAAACATATCGGGCACATTACGCATACGCAGCGTAAGCACGCAGGAGCAACAACTTTCGTCAGTGCCATTTGGCAGGGCTTTATCTAAAGCGGCGATTAATTGCAGTACATA
AATCGGGCTAATATGCAGAACAAACAAAAAAAAACTTCGGCTTCCCCCTAGCAGCTAACGTGGCCCCCAGTGGCGAATAAATCTCGTCGGCATCTTGCGGAATTAAAGAGAT
ACTTACGCACAAAGATGTTGCCAATCAGTGTGGCTGGCAGGCAGAAGACAATGACCAGGATGTCCGCAATGGCCTTTTCGTTGGAGAGGGGGGAATGTAAATTTTATTTCAA
ATTAATGGGGTTCATCCACGAGATTTACGGGGTCCACCCACCAGATTCACAATGAAGTAGTTCGTCACCGTCCGCATGCGAGGAGCCCGCAGGACGACGGCGATCACGAAGC
TATTCCCGATGAGTCCCACCAGGAAGACGACGATGTAGGCCACGCAGTAGACCATGGACATGGCCAGCGAGTGACGGTAGAGTGGGTCGAATCCGAAGTCCACGCTCCCGTC
CAGCTGCTCCCACCGGAGGGCACTGTCGTTGTACGACTTGGTGGCCGCAGCCGCCGAGATCAAGCTGGTGGTCGCCGGCGATAGGTTCAGCAAGTTGGAATAGTAGGTGGAT
GAAGTGGCGTCTGCGATGGCCAGCGACTTGCCCGCCAGGGATCCTCTCGCAGATGCTGTTGTGGTGGCCGCCAAGGTGGCTAGCAGCGAACTGGCCGCCGGAGCCACCGTGG
CCGTGGCCGTAGTGAGAACCGTGCCCAGCAGAGTGTCCGTAATGGCGCTGTTATTCAGTTTAATGGCACTCGATGAGCCCAAGGAGCTGACAGCAGTCCCACCGCTACCACT
AGTGCCATTCAATTCGTTGAAACTCAAGAGGTTCTGCAAATCGACGAAATTATGCGTGAAATTATTGTAATCCCCGGCGAGGGCGGCGGCTTCCTCCATTGTCTCCGCGACC
ATTTTCCCCGGCGCCATGGTCGGAATGGGCATGGGTACACTAGTCGTCGTACTGGGAACGTCGCCACTTGTATGGCTTTTATGCTTTCGCTTGCGAATCCGCCCGCTGGCTG
CCATCATCAAATCATCCGGCGTCTTGATTGGGCGGATGGCACTTTCACTGGCGCCTCGCTTTCGCTATCACTATCTGTAGACGTATCTCGAATCTTGAATCTGGTATCTCGC
TTGTTTTAGAAAAGGTCTATAAAAATGGCATCGTCACTTGGCATTTCGTTGGCTTGTGTAGGATATGCAGGAAGTTTCGTGGCTTTGGGTTAGAGAACAGTTTTTAATTAAC
ACCCATGTCCATACAATATCTCGTTTATCCGGGGATTGTTTTGCGTTTGCCCCGAAGTTGCACAATAAATGCAGTTTTCTAGAAACTCGCTTCTAGCGCCCAAGGACCCTTC
CGTTTTCGTTTTTTTTTTTTTGATTGAAATCACCTACAAAAGTTTGTTGCCGTTTCGGCTGCGAATCCGAGTGCGGAATGCAATGCATTGCCTGGGTCGTTCGTCAGCTGG
TTGACGTTGCCTTGGCTGGCATTACGAAAAAGTCGGCGGCCCAAGGCCTCGTGTGCATATTTCAGATAAGGCCAGATGCTGGGTGCCTGGATAAATGCAGCTAGAGGGGATT
GGGGATCGGGCATTGGGCATTGGGCATTGGGCATTGGGCATCGGACATCGGGCAAAGGAGCCGAGAGCATCGGCTCAGGGACATGGTCACCGCCACAGTTGCAACACATGCG
CGTCGCTAATTAATTTTAATGCCCGTTCTGCCATTGTTTGGCTGTTGTTTTTGGCACAATTGAAAAGCCAGCGAGCGTGCACTTTGACCTCAAAAAATCGACACCAAACATC
GAACAAACGAGCAATTGAGCAACTCAATAGCAGCGCATAAAAATTCCATTAAGCACAACTGCCAAGCGCACGCAGCACTCTTCGATTGAGTGCCAAAAAATTGCCAGTAAGC
ATTGAAACAATTGGTAACAAACGAAGCCAGTGCGACATTCTTTTGTGCGGTTCCATGGAGTGTCCACATCAGACAAACTGGCGACTAAGGGCCCCTCGGCACACAATTCA
(SEQ ID NO: 43)

Exon: 13111..12362
Exon: 12281..12156
Exon: 1649..1001
Start ATG: 13111 (Reverse strand: CAT)

Transcript No. : CT18916
ATGATGGCAGCCAGCGGGCGGATTCGCAAGCGAAAGCATAAAAGCCATACAAGTGGCGACGTTCCCAGTACGACGACTAGTGTACCCATGCCCATTCCGACCATGGCGCCGG
GGAAAATGGTCGCGGAGACAATGGAGGAAGCCGCCGCCCTCGCCGGGGATTACAATAATTTCACGCATAATTTCGTCGATTTGCAGAACCTCTTGAGTTTCAACGAATTGAA
TGGCACTAGTGGTAGCGGTGGGACTGCTGTCAGCTCCTTGGGCTCATCGAGTGCCATTAAACTGAATAACAGCGCCATTACGGACACTCTGCTGGGCACGGTTCTCACTACG
GCCACGGCCACGGTGGCTCCGGCGGCCAGTTCGCTGCTAGCCACCTTGGCGGCCACCACAACAGCATCTGCGAGAGGATCCCTGGCGGGCAAGTCGCTGGCCATCGCAGACG
CCACTTCATCCACCTACTATTCCAACTTGCTGAACCTATCGCCGGCGACCACCAGCTTGATCTCGGCGGCTGCGGCCACCAAGTCGTACAACGACAGTGCCCTCCGGTGGGA
GCAGCTGGACGGGAGCGTGGACTTCGGATTCGACCCACTCTACCGTCACTCGCTGGCCATGTCCATGGTCTACTGCGTGGCCTACATCGTCGTCTTCCTGGTGGGACTCATC
GGGAATAGCTTCGTGATCGCCGTCGTCCTGCGGGCTCCTCGCATGCGGACGGTGACGAACTACTTCATTGTGAATCTGGCCATTGCGGACATCCTGGTCATTGTCTTCTGCC
TGCCAGCCACACTGATTGGCAACATCTTTGTGCGTAAGTATCTCTTTAATTCCGCAAGATGCCGACGAGATTTATTCGCCACTGGGGGCCACGTTCATTGCCATCTGGTGGC
CCCTGAAGCAGATGACAAAGCGACGTGCCCGCATCATGATAATCGGCATATGGGTAATCGCGCTGGTGACCACCATTCCTTGGCTGCTCTTCTTCGACCTGGTGCCCGCCGA
GGAGGTCTTCTCCGACGCCCTGGTCTCGGCCTACTCGCAGCCGCAGTTCCTCTGCCAGGAGGTGTGGCCCCCGGGCACGGATGGGAACCTCTACTTCCTGCTAGCCAACCTG
GTGGCCTGCTACCTGCTGCCCATGTCCCTGATCACGCTCTGCTACGTGCTCATCTGGATCAAGGTCTCCACGAGGTCCATCCCCGGCGAGTCCAAGGACGCGCAAATGGACC
GCATGCAGCAGAAGAGCAAGGTGAAGGTCATCAAGATGTTGGTGGCCGTGGTCATCCTGTTTGTTTTGTCCTGGCTGCCGCTCTATGTCATCTTTGCGCGGATCAAATTCGG
CTCGGACATCTCGCAGGAGGAGTTCGAGATCCTCAAGAAGGTGATGCCGGTGGCCCAGTGGCTGGGCTCCTCGAACAGCTGCATCAACCCCATACTCTACTCAGTGAACAAG
AAGTACCGACGCGGATTCGCGGCCATCATCAAGTCGCGCAGCTGCTGCGGACGCCTCAGGTTAGTTAAT
(SEQ ID NO: 44)

Start ATG: 1 (Reverse strand: CAT)

MMAASGRIRKRKHKSHTSGDVPSTTTSVPMPIPTMAPGKMVAETMEEAAALAGDYNNFTHNFVDLQNLLSFNELNGTSGSGGTAVSSLGSSSAIKLNNSAITDTLLGTVLTT
ATATVAPAASSLLATLAATTTASARGSLAGKSLAIADATSSTYYSNLLNLSPATTSLISAAAATKSYNDSALRWEQLDGSVDFGFDPLYRHSLAMSMVYCVAYIVVFLVGLI
GNSFVIAVVLRAPRMRTVTNYFIVNLAIADILVIVFCLPATLIGNIFVRKYLFNSARCRRDLFATGGHVHCHLVAPEADDKATCPHHDNRHMGNRAGDHHSLAALLRPGARR
GGLLRRPGLGLLAAAVPLPGGVAPGHGWEPLLPASQPGGLLPAAHVPDHALLRAHLDQGLHEVHPRRVQGRANGPHAAEEQGEGHQDVGGRGHPVCFVLAAALCHLCADQIR
LGHLAGGVRDPQEGDAGGPVAGLLEQLHQPHTLLSEQEVPTRIRGHHQVAQLLRTPQVS*
(SEQ ID NO: 45)

Name: Tachykinin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384554
TGTTTGCAGAGCCCACAGCAACCAATATGTTTTCATTGTCTGCCCGGTTTGTCGCTCAATTTATTTGTGATGTTGACTTTCCAACAGTGACGTCAGTGGAGGCCACAGACAA
TACCACATCATCGCAACAATTATAAATGGTGTGGTGGCGAAAAAACGTCCTTGTGCCCATTGTTCAGTGGGGCGAACTGAAAAGTGGATTTTCCAGCGCGTGCCAGGAAAAT
GAATGAAAATGCGGGACATTCTTGAAGTCGTTGCCTTTATTCTGCTTTCACACATAATTATGAAATAAATATTAGCTTAAACCAAATTGACTACAACGTTATGCTCTATTAA
CTCTTATGGTCTACTTATGCTTAATGGGTTTTGGTTTAAGATGCTTGGGCATGTTTATCCACCATTACGAGTGGCTATATACACGGATGAGTTGGAATTTAAATGGCCTTAA
TAGCTAGTTAGGACTGGAAAGAAATGCATTTTAAATTTCTGATAAATGGCCGCGATTCACATACAAGTAAAATCAGTCTGTTTAAGAGAATATAACTAAACCGAGCTACTAA
TGGATTGAAACCTAGGTTCTGAGTGGTGGACAAAAAATCTTAATAGTTTTCGAAACACTTATTTCATTAGGATACCATATACTTAAATTAATTTCTCCAGTTTAAAGTTGAT
TGCAAAAAATGCCTAATTAAATGTTTCAGGTAAGTTTTTTGGAAATATAAAACCTTTGTACCAATATCAAATTATAATGGGTTCTTTAGTTAAAATTTAATAAAAATCAAAAT
CTTGAATAGATTCAATTACCCATTTTGACACTTTAGATTTTTAATTTTAGAGTTAATTAAATGCCGTATAAATGTTATCTTAACTATAGATTGCGAAGTAGAAATAGGTTTG
AAATTCGTAGCGTAGACTTGGCGTACTTACGACTAGAAACATACATTTGGCTTTACCCATGCCTGGTATAACTATCGAACTTATTGTCTAAGCCTCTGGTGCCTTCACAGGA
TGTCCGTGTTCCCCTTTAGAGAAACTCCGCGTCGCTAAAGGACGTGGGACCATGAATCGAGAACGCGGATCTTTTCCGGCTGGTGTGGGCCAGCTATCATACCTCCGACATG
GCTGGCACGGTAACCACCTTGGTGGGACGTTCGCAGATTACCACACGGGCGTTCGTCTTGTTGGTGGACCTTCGGGAAACCGTCAACGAAGTGGTCAGTGTCCTCATCGAAT
CGCAACTGTTGTGAAGCCGACGACCAACCGTGTGGCAGCGGTTTTGCTGCGAGTTGTTGCGGTATGACTTGCAGCAGCATGTGAACCACTTAAAAGGCGGAAAGCGACTGGA
AATATAATAAAAATAAAAATATCAAAGTAAGTAAATTAAAGCATAAGAAAATGTGTAAGATTTTTTAACATGAAATCTTAACATACCTTAATGTGCGAAAGACCTGCGATGA
GAAGAGGCAATAGATTAGTGGATTCGCCGCCGAGTTCAGCGGTGCCAGACTTTGGATGAAGGTGGCGATGGCAATGTTGGTCTGTGAGTGTGGAATCTGGCCAAAGACCTGC
AGCAGATCGAAGATGATATACGGCGACCAGCAGATGATGAACACAAACACGATGGTCAATGTCATCTTGACCGTTTTGACCTTTGCCCGTGGAATAATGCCCCTCGAGCTGG
CCCTCCTGGCAGGTGCAGCTCCAAAACCAGCACGTTCTGTTAAAATATAATTAACAAATAAAATAAAATTATTTTTGGGAACTCCGATTTGGCAATATGCATAGCCAGAAGA

TATCTTACCCGTGGGTACAAAAATGGAACCCTTTGCCCAAATCGTCTTTACGATGATCGCATAGCAGGCAGATATGATCAGCGCAGGAATGGCAAATAGAGTGGCCGACACC
AGGCTCATGTACACCTGCCAGGCGATCGGTGAACCCAACTCAATCCAGCATTGCGGATGTCCTTGGATGAGCTTCTCCTCGTACAAAACCAGGATGGGAAGCGAAAACAACG
CCGAGATGAGCCATGCGCCAGCCACCAGGTGACGGGCTCTTTTCCCTAAAAGCATAAAATTTAAATATATATTTAAGTGAAATATATTTCTAATTTAACTTAAATTTGTATA
TTTTTTTTGGTTAATTCCCGAACTAAATTATCCAAGCAGTGTAAGCAGAAGCCACAAACAGCGCTCATCTTTGGCTAACAACATTTGGCAATTAGAAAACTTTTCATACGAC
GACTTAACTTACACGACTTTGAGAAGTTCATGGGGTGTGTGATGGCATCGTATCTGTCGATGCTCATGGCCACCAGCACGTAGGTGGACGAGTATGTGACGCAGACCTGCGA
GAAGCGGATGGCCTTGCAGGCCAGGTTGCCTGCCCGCCACGAAATCGTGATGCGCCATATGATGTCGGTGAGGACGTTGAGCAGTCCCACGCACAGATCTGGAAATAATGGA
CTGGGCATTAGAAGCAAGGGAAGAGGTCAGTCCCAATCCCAAAACTCACCTGCCAATGCCAGCTGTTTAATGAAGTAGTTCATCCGCGACTTGCGATTCTTGTTGATGAACA
TCACGAACAGAACAGCTGAATTGCCCAGAACGATGACGGTGAACAGGATCCAGAGCACAGCAAACTGTTCGGTCTGGAAGAGTGCAAAGAACAAAGTGTGGTCACATTTCAT
TAAACGCCTATCAAGTTCATTGTCAACAAAGTGTTAAGCACGGAAATCACGCATAAAACTCAATTTAATGCCATCCAAGCACACACACACACACACGCCTTGTTTATGGCTC
TTTAAATAGCCATAAAAATGATAGAGCTATGAATTCAAGCTGGTTGGTGCGGCGGGCAGGATGTGCCGGCGGAAGTTTGCTTCTTCCAGTTCAGTACTAACAAGACTAGGAC
AACACGACATCCTGCACTCAAACACCCGTGGACGCTCACCCACACACCCAGACGTGATGGAAGACAGACAGCCACCGTTATTCCAGCAGGAAATTGCTGCCAAGATGAAGGA
CCTGCACTTCGAAAAATATTTCTTTCGACTTTTAAGGCAGTTTCACAGCTCAAAATGCGTGGCTTGGTATTTAATTTATAGAACTTATACTCCAGGCATTCGAAATTCAATT
CATAGCACTCAAATCTAAAATTGTAACAATTCTTATATAGAAAACTAGTAGATAAAGAATAATGAACCCTTATTACTTTTACTTTTTTAACTGAACTGTTATTATTTTAAAA
ATATTTTAAACAGTGTGAAGTGCAGATATGACAGTGAGTAGTTCTGCTAGTTCTAGCTTATTGCCGAAGAAAATGAAATATTCTTTATACGCCCGCATGCAGTTTATATAGC
AGAATATCGCATATATGCTAATATGACTAATGATCTCTTCATGGTATCTACAAAGTTTTTTGCTGCAAATTAAATTTGCTCTCAAGGCATCTCCCCCTATGCCCACGTCTAT
TGAAGTTAAGTTTGAATTATGGCCTCGGGGTCTTGCAGTCTGACGCCAACACCGATGTTGATGCGGACTTACTTAAGGGCAGGCCAAGTAAACTTGTCCCGAAAAAAAAGCT
GAAATCCTGACAGGCTTCGTCCTTGGAACTGCGCTCAGACTGAAGACGCCCCACCTTATTCTACAACTTTTTGTCCATACCAAACTTGCAGTATTCACTTTTTG
(SEQ ID NO: 46)

Exon: 2688..2514
Exon: 2450..2253
Exon: 2061..1801
Exon: 1716..1431
Exon: 1339..1111
Exon: 1042..1001
Start ATG: 2688 (Reverse strand: CAT)

Transcript No. : CT19191
ATGAAATGTGACCACACTTTGTTCTTTGCACTCTTCCAGACCGAACAGTTTGCTGTGCTCTGGATCCTGTTCACCGTCATCGTTCTGGGCAATTCAGCTGTTCTGTTCGTGA
TGTTCATCAACAAGAATCGCAAGTCGCGGATGAACTACTTCATTAAACAGCTGGCATTGGCAGATCTGTGCGTGGGACTGCTCAACGTCCTCACCGACATCATATGGCGCAT
CACGATTTCGTGGCGGGCAGGCAACCTGGCCTGCAAGGCCATCCGCTTCTCGCAGGTCTGCGTCACATACTCGTCCACCTACGTGCTGGTGGCCATGAGCATCGACAGATAC
GATGCCATCACACACCCCATGAACTTCTCAAAGTCGTGGAAAAGAGCCCGTCACCTGGTGGCTGGCGCATGGCTCATCTCGGCGTTGTTTCGCTTCCCATCCTGGTTTTGT
ACGAGGAGAAGCTCATCCAAGGACATCCGCAATGCTGGATTGAGTTGGGTTCACCGATCGCCTGGCAGGTGTACATGAGCCTGGTGTCGGCCACTCTATTTGCCATTCCTGC
GCTGATCATATCTGCCTGCTATGCGATCATCGTAAAGACGATTTGGGCAAAGGGTTCCATTTTTGTACCCACGGAACGTGCTGGTTTTGGAGCTGCACCTGCCAGGAGGGCC
AGCTCGAGGGGCATTATTCCACGGGCAAAGGTCAAAACGGTCAAGATGACATTGACCATCGTGTTTGTGTTCATCATCTGCTGGTCGCCGTATATCATCTTCGATCTGCTGC
AGGTCTTTGGCCAGATTCCACACTCACAGACCAACATTGCCATCGCCACCTTCATCCAAAGTCTGGCACCGCTGAACTCGGCGGCGAATCCACTAATCTATTGCCTCTTCTC
ATCGCAGGTCTTTCGCACATTAAGTCGCTTTCCGCCTTTTAAGTGGTTCACATGCTGCTGCAAGTCATACCGCAACAACTCGCAGCAAAACCGCTGCCACACGGTTGGTCGT
CGGCTTCACAACAGTTGCGATTCGATGAGGACACTGACCACTTCGTTGACGGTTTCCCGAAGGTCCACCAACAAGACGAACGCCCGTGTGGTAATCTGCGAACGTCCCACCA
AGGTGGTTACCGTGCCAGCCATGTCGGAGCGACGCGGAGTTTCTCTAAAGGGGAACACGGACATCCTGTGA
(SEQ ID NO: 47)

Start ATG: 1 (Reverse strand: CAT)

MKCDHTLFFALFQTEQFAVLWILFTVIVLGNSAVLFVMFINKNRKSRMNYFIKQLALADLCVGLLNVLTDIIWRITISWRAGNLACKAIRFSQVCVTYSSTYVLVAMSIDRY
DAITHPMNFSKSWKRARHLVAGAWLISALFSLPILVLYEEKLIQGHPQCWIELGSPIAWQVYMSLVSATLFAIPALIISACYAIIVKTIWAKGSIFVPTERAGFGAAPARRA
SSRGIIPRAKVKTVKMTLTIVFVFIICWSPYIIFDLLQVFGQIPHSQTNIAIATFIQSLAPLNSAANPLIYCLFSSQVFRTLSRFPPFKWFTCCCKSYRNNSQQNRCHTVGR
RLHNSCDSMRTLTTSLTVSRRSTNKTNARVVICERPTKVVTVPAMSERRGVSLKGNTDIL*
(SEQ ID NO: 48)

Name: Vasopressin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384294
TTCCAATGCGATTTAAGGTAATTGACTATAAGTACCCATGACGAACTTCTGGACAGTTCCCCAGTGGAAAAAAAAAAACTGGATTTATTTCTTATCGCAAAACTTGCGATAAC
CAATTTGTATATGTTCATTCAGGGAGCGTGCCGATTTTTTAGTTTGCGAAATTGAGTTGTTATGTTTTCTGAGTCATTGTCGAGCGGGTTTTGGTAGCCGGGGATTAAATCT
GCTGTATCATCTTGGTCCAGCGTTTATTTCTATTTGCTTTTAATTTCACAATCAAAACAGCAATAATAAAACCGTGCCTTATCTTACGGACATATCTCCCCAAGCGTCGTCT
CGCTCTAATTTGGATTTTTCCTCTCTCTCTCGCTCTTTTCTTGCAGCTGGCGCCCGACGGTGGGAATTTCATCAGCTGTTGTTGTTGTTCATTTCACAGCAGCAGTCGGCCT
AATAAAAATGCAATTTTTTATGGTTAAGAAAATAGTGCGCGAACGCCGTTGTGTTGCACATTGGAAATTGTGCAGAAGCCAGAAGAAAATGTGTCTTGGAAAATGGAAAAGCA
CGAGCAGCAGTAGCACCAGCAGCAGCTATAACATTGAGCCAGTGATGTAATACAGAGCAAAAGTGAAGCGGCGGCAATTGGCATTGCAAAAAAGAAACAGATTATACAATTT
GTGTTGCTGCCGCACTGAACTAGGACAACAGCAGCAACAACAGCAGCAGCAACAACAACCACTAGTGATAAAGACCGGCTGGCAACAGTGCAGCGGCCACTTGCAGCAACAT
AATTTCAAAGTCGGAGCGTGGGAATCGCAAATCACGGCAGTGGCAGCAACAACAGCAACTGCAGCAACACTCGAGCACTCGTGGCTCTCCGGATCGCTCATTTCGCTCGCCG
TTCTCGCGCTGGCGTTTCTCCGTTTCGCTCGCACCCGCCTGTGGAGCAGTCGCGTATATTTGGAGGCGTTGGGGCGAACGCTCACGACAGTTGTGCTTGGAACGCTTCGGAG
TGAAGATCGTGCTCGTCCGATCGGTGCGCGTGATAAATACACGGCGAACAAAATTGAGTCAAGAGTCAGACGGTCGCAAGACAATTGACAAGACACGAAAAATCATAATAAA
CACAAGTGGAAAAGCAACTGCTTGAGAGCAAAAAAGAAATTGTTTTTTACAAATCTCGCACGTAAATTAGCAAGCGTTCGCAAGAAACAAACCGCATACAATACTGAACACA
AAAAAAAGGGTATTCAACAAAAAACCAAAAATATTTGGCTGCAAATGTGAAGAGACGAGAGACAGCGGATATACTACCGAGTGAAATTTGTGTCATAAGCAGAGCCAAAACA
AAACCGAAAAGCCAAATTCAATAAAAGGCCATAAGTGACTGTGAAACTGTAAAGTGATGTGTGGTAATCGAAAAACAGCATAAAGAATGGCCATAGACCTCCTGATCCTGGC
CCTGCTGCTCGTCTCGTTCCTGATCAACCTGCTGGCGCTGGTGCCTTTTGGATAACGCCTGGCTTGAGGACCACCGCCAATCGCTTTACGATCAACCTGCTGGCCATCAAT
CTCATTGGTTGCTGCATCTTGGCCCCCACTTTGTTCCTGGGACTGCCGGGAAAGTCTGCGGAGGCATCCACTTCCAACGCCGAGACCCTGGAGTTCTTCTCGAAGCCGGGAA
ACCATCAGGTGATACTTCGTCGCAATGGCCAGCTGGTGGAACAGGACGGCGTGGTGGTGCGAAGGAACATCAGCGAAAACGGCGATACGGTGGAGACCTTCTTCAAGTGCAA
TGCCACCTACTGTCGCGAGTTGACCATCGATGAGCGCGGAGATGGTGGCTTTGTCATCACGGAAACCGAGACACACGAAGAGAATCTCTCGGCCTTTGAGAGCTTACCCACG
GAGGCTCCCATTTTGCCGCCCGTCCAACTGAGATGTTGGTCCATAGACATGACTGCTGCGCTGGGAGCTCTGGCAGTACTGCTCGTGGTGGGGGACACCTGGTGTGCCGTTA
CGGATCCCCTGCGCTACCACAGTCGCATTTCGGGCGTGAAAACCTGGATCTTCATAGCACTCACCTGGGTGGTGGGCATACTCTTTGGAGCCCTTTCCGCCTTCCGAGTTTT
GGACTTTGAGGCGGACGCCCTGTTCAGCAGACAACGCCGCCTGGCAGTCACGTACTTCAACATCAGCAGCACCAATAGTATTTTCGGAGTGGTGTACGCCAGTGTCTACTTC
ATCGTGATCATCCTCCTGCCCTTCGGCTTCGTGTGCGGAATGTACTGGAGGATATTCAGTGAGGCGCGCGGAAATGGGCTGCGCATGCGCCAGAATGGCTCTTCGCCGCTGC
TGCAGAGCGCCTTGAATCTGACGGCTGGCCAGCAGGCGGCTCAAGCCAACCAGTTCTCCAATAGCTTGTGTGTTCATCGGCATTCCATATCCTCGGCCAGTTCGCATGGCGG

CAACAGCAGCTTGGGCTTGGGCGGTCTGCAAATGCAAATTGATCAGCGCCAACAGCCGAGGAGTTCGCCCAGTTGTCTCCGTCGGGACTCGGCCGCGAAAGTACTTTTGCCC
ACCATCTCGGATGATGGTGGCTCAGATGCGGAAAGTGGAGCCGGAGTTCAACTGATGCCCGTACAGGAGCACTCACTGTCAGATAGAAATCAGAACATTATGCTTACTCTGC
AAACAGCCAGTGGTGAAATAAAGAGGAACTACTCCGCCAGACAGCTGCCCCTGTTGGGCACTTCTTCGCAGGATCTGAGGGAGACTAATCGTCTGCAGGGCATACGCCAGGT
GCACAGTTCGCCCAATCTCCACAAGTACACGGAACTGCGCCAGGACTCGCTCAGTGAGGAGTGTGGTTCTCCGCATCTTTTGGGTCACGCCCAGCGGCAGCAGCAGCAGCAA
TTGCACCTGCAACATCAGCAGCAACAGCATCATCACCACCAGCAACATCATCCGCACTTCAGTTCGCCGCGCCACCAGCAACACGGCCACGCCCTCCAGATACCCGCCATTC
ATGCCTCCCCCAAGGCTCTGAGTTACATGAGCTCCTTGCGACACCGTTTGAGCAACGCCAGTTCGCTATTCAAATACCGTGAGGAGTCGAGAGCAGCCCGCATCAGCATCCT
GGTGGTGGTGATGTTTGTGGTCTCGTACCTGCCCTTCGGACTGCTGGTACTCCTGCAATCGCGACTGTCGGCTGCTAACTTTGGCGGATCCTCCCAGCTGGCCATCTTCATG
ATCCTGCTGGCCAATCTCAGTTCGCCATTCATCTTCGCCTACAGGAACAAGAGGGTGAGGCGAGGAGTCAAGCGGCTGTTTGGTCTGGATTCGTCGTCGGGCCTGCAGCGCA
ACTGCAGCAGCAGTGTAAAGACCAATGGCACCGCAGGTCCTGCTGCAAGTGGTGCACAATTGCAGCGCAACAGCAGCAAGTTGTCGCAGTACAGCAGCAACAGCTGCAAGTA
TCTCACCCCCCAGAGTTCCCTGGTCAGTCAAGTGCCGGTGCATACGACCTTGACGTTAAGGCCCAACAGCAGCTGCAGCACCATCATCAATTTCGGTGGCGCCAGGGGCTCA
GCGGACTCCGATGAGCAGCCACCGGCAACACCACCACCTACTGTGGCAGTTGCTCCACCGCCCACACGACCGCAGCGCCCCAAGCTGCAGAGGGGCATCACCATCGTTGAGC
ACATAGCTATAACACCAACCATGCCCCAAAAGTTCCAGAATCGGCGCGCCAGGCTCTTCGACATGTTCTTCCGCAGCTCCAAGAAACTGCAGGCGGGCTGCCAGAGCCAGAG
TTTGCCCACTGAGGTCTAAACGGGGGGTTATGATAACTATACATATAGGAAGGTAAACAAACACTAAACTTGCTTGCTGAAGAATTCTAAGAACAGTATTCGTAATTTGTAA
GTCTAGCAATTAACAAAGACAACTGTTTGAATTTGAGTTTTAAGACCAGATCATGCTGTCCTAGAAGCTTCTCATGCCAAAAAACAAAACCACAAAGAAAGTCAAACTAGAA
AGTTTGAGATGCTTTAAATTAACAAAAAAAGAAAAGAAAAAAACAAAAAACAAAACAAAGTACCTTTTAAAAAATAGAAAACTGTGTACAAGTCAAATTAACAAAGATTGAG
AGTGATGGGAGTTATAATCTTAGGGTGGATATCAAACTGAAAAGTGGAACCAAGCCAAGCCAAGTTGAGATCGTAATCAATGTACCTAAGCAACTTGCATATGAACATTTAA
TGTTTTTACTCAGATGAAATCACAGGAAAATAGCAAAGACAGCACAGCTTATTCGAACTTCCCCAACTAGTTTATTATTTATTGAAACAATTGTAAATATTTATTAGACATC
CACGCCGCAAAGCTTTAACTCCAAAAACATTCTGGCTAAACAAGAAGAGACGCATACATTCGATTCATAAGTTAATTAAGTCAATGTGCAAAGCATATGTGTATTCTAGAGT
TTAAGTTGTAGAGCCGAGTAGTCTTACACCGAGCACCTAGTGCCGAGCAGTGCTTTTTCACCTGCCCAACTGGAGCGTCTGTTGCAGGCGGCATTTTCATATTGAAAATGCA
TTTTGGCCAGTGCACAGTTGGTTTTCCAGTTGGCCAGAATGTCAGTCAGCAGTTAGTTAGCCTTTGGTTAGCCAACGGCTATTGGGCCATTAAGTCAATTTGCTTTTTAGAA
CACCGCGGGCATTGCGTGCAGCTAAATACCTCATTAGTTATAATTTGAAGTGCGAGATGGATGGCGGCAATAGGTCAAAGGTCGGAGCCATCGGCACGGCCCATGCATAATG
CATGGAAAATGTTCGATTGCCCCTGGCATTTGTATCATTTGCTTTACTATGCTGCAGGCCAGAAAGCTTTATACCTTAAGTTCTCCTGCTTTGATTTCCCTTGGCTGCACAA
AAAGGCAACAAAAGCAACACACACTCACACAACACAAAAAAACCTCGAATTTATATTTCTAAAGTTTATATTTTATAAAGAGTGTGTGCATAGGAAATATCTATAAATATTC
ATGCATGTGTG
(SEQ ID NO: 49)

Exon: 1001..4051
Start ATG: 1431

Transcript No. : CT19320
CTTCGGAGTGAAGATCGTGCTCGTCCGATCGGTGCGCGTGATAAATACACGGCGAACAAAATTGAGTCAAGAGTCAGACGGTCGCAAGACAATTGACAAGACACGAAAAATC
ATAATAAACACAAGTGGAAAAGCAACTGCTTGAGAGCAAAAAAGAAATTGTTTTTACAAATCTCGCACGTAAATTAGCAAGCGTTCGCAAGAAACAAACCGCATACAATAC
TGAACACAAAAAAAAGGGTATTCAACAAAAAACCAAAAATATTTGGCTGCAAATGTGAAGAGACGAGAGACAGCGGATATACTACCGAGTGAAATTTGTGTCATAAGCAGAG
CCAAAACAAAACCGAAAAGCCAAATTCAATAAAAGGCCATAAGTGACTGTGAAACTGTAAAGTGATGTGTGGTAATCGAAAAACAGCATAAAGAATGGCCATAGACCTCCTG
ATCCTGGCCCTGCTGCTCGTCTCGTTCCTGATCAACCTGCTGGCGCTGTGTGCCTTTTGGATAACGCCTGGCTTGAGGACCACCGCCAATCGCTTTACGATCAACCTGCTGG
CCATCAATCTCATTGGTTGCTGCATCTTGGCCCCCACTTTGTTCCTGGGACTGCCGGGAAAGTCTGCGGAGGCATCCACTTCCAACGCCGAGACCCTGGAGTTCTTCTCGAA
GCCGGGAAACCATCAGGTGATACTTCGTCGCAATGGCCAGCTGGTGGAACAGGACGGCGTGGTGGTGCGAAGGAACATCAGCGAAAACGGCGATACGGTGGAGACCTTCTTC
AAGTGCAATGCCACCTACTGTCGCGAGTTGACCATCGATGAGCGCGGAGATGGTGGCTTTGTCATCACGGAAACCGAGACACACGAAGAGAATCTCTCGGCCTTTGAGAGCT
TACCCACGGAGGCTCCCATTTTGCCGCCCGTCCAACTGAGATGTTGGTCCATAGACATGACTGCTGCGCTGGGAGCTCTGGCAGTACTGCTCGTGGTGGGGGACACCTGGTG
TGCCGTTACGGATCCCCTGCGCTACCACAGTCGCATTTCGGGCGTGAAAACCTGGATCTTCATAGCACTCACCTGGGTGGTGGGCATACTCTTTGGAGCCCTTTCCGCCTTC
CGAGTTTTGGACTTTGAGGCGGACGCCCTGTTCAGCAGACAACGCCGCCTGGCAGTCACGTACTTCAACATCAGCAGCACCAATAGTATTTTCGGAGTGGTGTACGCCAGTG
TCTACTTCATCGTGATCATCCTCCTGCCCTTCGGCTTCGTGTGCGGAATGTACTGGAGGATATTCAGTGAGGCGCGCGGAAATGGGCTGCGCATGCGCCAGAATGGCTCTTC
GCCGCTGCTGCAGAGCGCCTTGAATCTGACGGCTGGCCAGCAGGCGGCTCAAGCCAACCAGTTCTCCAATAGCTTGTGTGTTCATCGGCATTCCATATCCTCGGCCAGTTCG
CATGGCGGCAACAGCAGCTTGGGCTTGGGCGGTCTGCAAATGCAAATTGATCAGCGCCAACAGCCGAGGAGTTCGCCCAGTTGTCTCCGTCGGGACTCGGCCGCGAAAGTAC
TTTTGCCCACCATCTCGGATGATGGTGGCTCAGATGCGGAAAGTGGAGCCGGAGTTCAACTGATGCCCGTACAGGAGCACTCACTGTCAGATAGAAATCAGAACATTATGCT
TACTCTGCAAACAGCCAGTGGTGAAATAAAGAGGAACTACTCCGCCAGACAGCTGCCCCTGTTGGGCACTTCTTCGCAGGATCTGAGGGAGACTAATCGTCTGCAGGGCATA
CGCCAGGTGCACAGTTCGCCCAATCTCCACAAGTACACGGAACTGCGCCAGGACTCGCTCAGTGAGGAGTGTGGTTCTCCGCATCTTTTGGGTCACGCCCAGCGGCAGCAGC
AGCAGCAATTGCACCTGCAACATCAGCAGCAACAGCATCATCACCACCAGCAACATCATCCGCACTTCAGTTCGCCGCGCCACCAGCAACACGGCCACGCCCTCCAGATACC
CGCCATTCATGCCTCCCCCAAGGCTCTGAGTTACATGAGCTCCTTGCGACACCGTTTGAGCAACGCCAGTTCGCTATTCAAATACCGTGAGGAGTCGAGAGCAGCCCGCATC
AGCATCCTGGTGGTGGTGATGTTTGTGGTCTCGTACCTGCCCTTCGGACTGCTGGTACTCCTGCAATCGCGACTGTCGGCTGCTAACTTTGGCGGATCCTCCCAGCTGGCCA
TCTTCATGATCCTGCTGGCCAATCTCAGTTCGCCATTCATCTTCGCCTACAGGAACAAGAGGGTGAGGCGAGGAGTCAAGCGGCTGTTTGGTCTGGATTCGTCGTCGGGCCT
GCAGCGCAACTGCAGCAGCAGTGTAAAGACCAATGGCACCGCAGGTCCTGCTGCAAGTGGTGCACAATTGCAGCGCAACAGCAGCAAGTTGTCGCAGTACAGCAGCAACAGC
TGCAAGTATCTCACCCCCCAGAGTTCCCTGGTCAGTCAAGTGCCGGTGCATACGACCTTGACGTTAAGGCCCAACAGCAGCTGCAGCACCATCATCAATTTCGGTGGCGCCA
GGGGCTCAGCGGACTCCGATGAGCAGCCACCGGCAACACCACCACCTACTGTGGCAGTTGCTCCACCGCCCACACGACCGCAGCGCCCCAAGCTGCAGAGGGGCATCACCAT
CGTTGAGCACATAGCTATAACACCAACCATGCCCCAAAAGTTCCAGAATCGGCGCGCCAGGCTCTTCGACATGTTCTTCCGCAGCTCCAAGAAACTGCAGGCGGGCTGCCAG
AGCCAGAGTTTGCCCACTGAGGTCTAAACGGGGGGTTATGATAACTATACATATAGGAAGGTAAACAAACACTAAACTTGCTTGCTGAAGAATTCTAAGAACAGTATTCGTA
ATTTGTAAGTCTAGCAATTAACAAAGACAACTGTTTGAATTTGAGTTTTAAGACCAGATCATGCTGTCCTAGAAGCTTCTCATGCCAAAAAACAAAACCACAAAGAAAGTCA
AACTAGAAAGTTTGAGATGCTTTAAAT
(SEQ ID NO: 50)

Start ATG: 431

MAIDLLILALLLVSFLINLLALCAFWITPGLRTTANRFTINLLAINLIGCCILAPTLFLGLPGKSAEASTSNAETLEFFSKPGNHQVILRRNGQLVEQDGVVVRRNISENGD
TVETFFKCNATYCRELTIDERGDGGFVITETETHEENLSAFESLPTEAPILPPVQLRCWSIDMTAALGALAVLLVVGDTWCAVTDPLRYHSRISGVKTWIFIALTWVVGILF
GALSAFRVLDFEADALFSRQRRLAVTYFNISSTNSIFGVVYASVYFIVIILLPFGFVCGMYWRIFSEARGNGLRMRQNGSSPLLQSALNLTAGQQAAQANQFSNSLCVHRHS
ISSASSHGGNSSLGLGGLQMQIDQRQQPRSSPSCLRRDSAAKVLLPTISDDGGSDAESGAGVQLMPVQEHSLSDRNQNIMLTLQTASGEIKRNYSARQLPLLGTSSQDLRET
NRLQGIRQVHSSPNLHKYTELRQDSLSEECGSPHLLGHAQRQQQQQLHLQHQQQHHHHQQHHPHFSSPRHQQHGHALQIPAIHASPKALSYMSSLRHRLSNASSLFKYREE
SRAARISILVVVMFVVSYLPFGLLVLLQSRLSAANFGGSSQLAIFMILLANLSSPFIFAYRNKRVRRGVKRLFGLDSSSGLQRNCSSSVKTNGTAGPAASGAQLQRNSSKLS
QYSSNSCKYLTPQSSLVSQVPVHTTLTLRPNSSCSTIINFGGARGSADSDEQPPATPPPTVAVAPPPTRPQRPKLQRGITIVEHIAITPTMPQKFQNRRARLFDMFFRSSKK
LQAGCQSQSLPTEV*
(SEQ ID NO: 51)

Name: G-protein coupled receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384666
ATTCTATTGTCGTAGGCGTGGCAAAAGGTTTTTGGTAAATCGATAGAAATGTACAAGAGAATGTAATTAATATGAAAAGGTAACAAAAAATTTTGCAAAAGTGTGGGCATG
GCAGTTTTGGTCGGTTTGTTGGCGTAAGGGCATTATCGGTATTTTCCTATTTTCAGCTATATGATGCAGTCGCCTAATTTTGATCAAACTATAACTAAAATTGTATATTGTC
AGAAACAACGAGCATGAATTTATAGTCAATTTATCTTTTTATTATTTCTTATCTAAAGTTATATGCATATAGAATAAAAACAGATACGAGCTTCCTTAGAAGCGGACGGAAA
TAATGTCTCGAGATCTGGACCTTCATTAAGACAGACGAACATTAGGACAAGGCCAGGTCTACTCAGCTAGGGATCCTGATCAATGTATATGTATTTACTTTAAAGGGGCGTA
AACTTGTATTTTACATGTTACATACTTTCAAGCGAATCTAGTAAACCCTTTTAATTTACAAGTAACGAATATAAAAAGTTGAATGGTAGCGGGAGAAAATTAATTGTATAAT
ATAATATTATCTTTGTTAATCCAGAGTCTTGGCAACTTGTGCACACAATTTTTTTTTTTTTTTGGAAACTGAATCAGTTAATTGTGCCAAAATTTAGAGTGTAGATTTCTT
GGAAATCTCAATGAATTTTATGGTCAAGTTCCAAGAAGTGAAATGCTTAACAAACAACAAAAACAAAACACACAAAAGCAAACAAATTATGCAGAATTATCGGATGTATTTT
ATATGTATGTATGTACATACATATACAAAGAATAATAAATGGGTGGGTCTGTTGGCTTCTATGGAATTTCCAATTTAAGCTGCAATACATGTACATTTTAACGTCTTTCTGT
ACGCTCTCAAATAAAAAAACACAATTTTAGGATTAATAAAAAAACAGATATGTACGTCTTCTATTCCCATATGAAGAAAAATATTCATATGATTTTTTGTTTGGCAGGAAAT
GCGAATAGTTATTGGATCGTTCACCGCATTTCTTTTGCTGTTATTGCAAAACTCAAATGCCGAAATTCCCGGTTGCGACTTCTTCGACACCGTAGATATTTCAAAAGCGCCA
AGATTCTCGAACGGATCGTACCTCTACGAAGGCTTGCTGATCCCCGCCCATTTGACAGCTGAATATGACTACAAGCTCCTGGCCGACGATTCGAAGGAGAAGGTGGCGAGCC
ACGTACGAGGATGTGCCTGCCACCTCAGGCCATGCATTCGGTTTTGTTGCCCCCAGTACCAAAAGATGCAAAAGAGCAAGTGCTACGGCGACATGTCGGAGGACGAGCTGAA
CAAGCACGATCCCTTCGTGAACGTGACGCTCAGCGACGGGTCGGTGGTCAGGAGACACTTCAAGGAGGATCTGATCGTGCAGTCGGATCTGGCCAAGCCCGGATGTCCCCGG
ATGTACTTTCTGAATCACGAACTGCCGGGCAATGAATTCACTCTGTTTGAGGTTACTTTCCTGTGTTTCTGTCACTTCTGAAGATCGTAACCTTTTTTGAATCTCTGTAGAA
CGGCTCTCTTCTGCGTCACTGGGACAAAGTGGAACTGAGCAAGCCGGGAGTACTGCGTCCAGCATCTTTCATTTAAGGATGATAGTATCCGAATTGCACCCCACTTCTGTCC
ACTTTCGTCTGAGCACTCCAGAACGTGGAAGACCGTTGGTAAGCTGAATTCCCTGCTTGACTTTAAGGTTTTTAGTTAATGTCCATTTTCACAGCGATAGTGATATCCTTGA
TATGCATAATCCTAACGATCAGCGTGTACCTCTATGTTGAGAAGCTTCGCAACCTGCACGGAAAGTGTTTCATCTGCTACTTGGCCTCCTTGTTCCTGGGATATTTCTTCCT
GGTCCTCAACGTGTGGAAATACTCCTCCGGCTTCTGCGTTACAGCAGGTTTGCTTGTGGCGATTACAAGATAAAGGACGACCAATTTCACATTTCGATATCATGATCAGGGT
TCCTGGGCTACTTCTCCGTCATCGCCGCGTTCTTTTGGCTTTCTGTCATCAGTCTTACCCTCTGGAATTCATTTAGCGGCAACTCCAGCTGGTTGAACCGTTTCTTGCCACA
GAATCGGTTCCTTTCTTACAATTTATACGCCTGGGGCATGGCGCTGCTCCTGACTGCAATCACCTATATAGCTGACCAGGTGGTTAAGAACGAGAAGTTGAGACCCCGCGTG
GGCGTCGGCAAAAATTGTTGGATCTACAGTAAGTTACTATTCCTATAATAACAATTGCGGTGTAACTCGTTTTTTTTTTAACCAACAGCTGGGGATATGACTGTCATGATCT
ACTTCTATGGCCCGATGCTACTGATAATTGTTTTCAACATAACAATGTTTGTCCTGACGGCTTTTCGTATAATGAAAGTGAAGAAGGAAGCACAAAACTTTACTCAACAGCA
AAAAACAACCAATAGGCTCAATTCGGACAAGCAGACGTAATAATTTTACCATATTTTATACAGTTCTAAATCCAAAAGTTTAAATTTTAAATTATTTTCTATGGCTTTTAAA
GTTACGCCCTATTCCTGCGACTCTTTATCATCATGGGCTTGTCCTGGAGCTTGGAAATAATCTCGTTTTTGTTGAGCAAAAAATCAGGCTTGGGCCAAGGCTTTTATGGTGGC
TGATTACTTCAATTGGTCACAGGGCACCGTCATATTTCTGCTATTTGTTCTGAGGCCCAGCACACTGAAGCTACTGAAGGAACGGTAAGTTATCGGTTCTTCCATAGGAGGT
CCAATTAACCTGGTTCTATCATCTGGCTGAGCAGGATTAAGGGTGGGAGGGATGAAGCCGGCGCCAGCGATGAACATATCTCACTTCAGAATACGAAAATTGACCCCAGTGT
TTTTTAATATATGGGTAATGTGAAAATACAGTTTATTAGTTCATTTGGTGTTGGAAACCTTCTATTTTGAATCACTAACAGGCAGCCTAATGATCGTGACCAAAGGCTGCTT
TGCTTGCTCATTATTGCTCACATATCGTATGCTCATATGGGCTACGAATATTTGCTACTTATCACTTTTCAAGTATCTTGATTTCTGCACGGTGTCTAAGCATATCTAAGTT
TTAATTAAATGTTGAATGTAATTGTCCTTTAAGTTATAACTTTTTGCAATAATAATCTATTCCAATCCATCCTAGACGTGCCATCAAAGCGAGATCAAAGCAACTACTACAT
GGGGATATGTACGAATCTTATAGGCATGAATACTCGGCATTTATATACCCTGGTCTATTATTATGGATGGAGGATAAAAACTCAGGAAAATACTTTTCGCATCAGCACTTAT
GTATGCTCACTCGTGATAATAATTTCGTAATAATTCGTACTGTTTTTACTGCCTCAACTTAATTCAAAGAACTACAAATGATTATATTATGACTTTTTATGACAATGACTAA
CTAAAAATAAAAATTTCTGAATATTTATAAACAACTTTAAAATGTTATTAACTTAGGATAATAATAATAACACAGTTTTAATTTATCAAGCACATCGTTAAAAAACAATAAA
TTACTTGAATATAAATTTAAAAAAAAAATGGAATATTGGTTTACATTAACGGTATGCAAAAAATCATAAAAAGTCATAGAACTGAACGATAGAGAAAATTACAAAAAAAACGA
TCTAGTCGAGTTCCTCGACTATAAGATACTCATTTCTCAGCTAGTTCTTTCATTTAAAAATTTCTCTTTCGTTTCGATAGTAATTGGCAAGAAAAAATAATAAAATACAACA
AATTAATGCATTTTTGAAAATTATGGGCGTGACTGATTTGGTCGGTTAGTGGACGTGTCAATATTATGAAGCAAACGTGGACACTACAAACTGAAATATATATATATATAT
(SEQ ID NO: 52)

Exon: 1001..1507
Exon: 1567..1703
Exon: 1756..1951
Exon: 2014..2268
Exon: 2329..2500
Exon: 2578..2772
Exon: 2835..2919
Start ATG: 1007

Transcript No. : CT20339
CAGGAAATGCGAATAGTTATTGGATCGTTCACCGCATTTCTTTTGCTGTTATTGCAAAACTCAAATGCCGAAATTCCCGGTTGCGACTTCTTCGACACCGTAGATATTTCAA
AAGCGCCAAGATTCTCGAACGGATCGTACCTCTACGAAGGCTTGCTGATCCCCGCCCATTTGACAGCTGAATATGACTACAAGCTCCTGGCCGACGATTCGAAGGAGAAGGT
GGCGAGCCACGTACGAGGATGTGCCTGCCACCTCAGGCCATGCATTCGGTTTTGTTGCCCCCAGTACCAAAAGATGCAAAAGAGCAAGTGCTACGGCGACATGTCGGAGGAC
GAGCTGAACAAGCACGATCCCTTCGTGAACGTGACGCTCAGCGACGGGTCGGTGGTCAGGAGACACTTCAAGGAGGATCTGATCGTGCAGTCGGATCTGGCCAAGCCCGGAT
GTCCCCGGATGTACTTTCTGAATCACGAACTGCCGGGCAATGAATTCACTCTGTTTGAGAACGGCTCTCTTCTGCGTCACTGGGACAAAGTGGAACTGAGCAAGCCGGGAGT
ACTGCGTCCAGCATCTTTCATTTAAGGATGATAGTATCCGAATTGCACCCCACTTCTGTCCACTTTCGTCTGAGCACTCCAGAATTAATGTCCATTTTCACAGCGATAGTGA
TATCCTTGATATGCATAATCCTAACGATCAGCGTGTACCTCTATGTTGAGAAGCTTCGCAACCTGCACGGAAAGTGTTTCATCTGCTACTTGGCCTCCTTGTTCCTGGGATA
TTTCTTCCTGGTCCTCAACGTGTGGAAATACTCCTCCGGCTTCTGCGTTACAGCAGGGTTCCTGGGCTACTTCTCCGTCATCGCCGCGTTCTTTTGGCTTTCTGTCATCAGT
CTTACCCTCTGGAATTCATTTAGCGGCAACTCCAGCTGGTTGAACCGTTTCTTGCCACAGAATCGGTTCCTTTCTTACAATTTATACGCCTGGGGCATGGCGCTGCTCCTGA
CTGCAATCACCTATATAGCTGACCAGGTGGTTAAGAACGAGAAGTTGAGACCCCGCGTGGGCGTCGGCAAAAATTGTTGGATCTACACTGGGGATATGACTGTCATGATCTA
CTTCTATGGCCCGATGCTACTGATAATTGTTTTCAACATAACAATGTTTGTCCTGACGGCTTTTCGTATAATGAAAGTGAAGAAGGAAGCACAAAACTTTACTCAACAGCAA
AAAACAACCAATAGGCTCAATTCGGACAAGCAGACTTACGCCCTATTCCTGCGACTCTTTATCATCATGGGCTTGTCCTGGAGCTTGGAAATAATCTCGTTTTTGTTGAGCA
AAAATCAGGCTTGGGCCAAGGCTTTTATGGTGGCTGATTACTTCAATTGGTCACAGGGCACCGTCATATTTCTGCTATTTGTTCTGAGGCCCAGCACACTGAAGCTACTGAA
GGAACGGATTAAGGGTGGGAGGGATGAAGCCGGCGCCAGCGATGAACATATCTCACTTCAGAATACGAAAATTGACCCCAGTGTTTTTTAA
(SEQ ID NO: 53)

Start ATG: 7

MRIVIGSFTAFLLLLLQNSNAEIPGCDFFDTVDISKAPRFSNGSYLYEGLLIPAHLTAEYDYKLLADDSKEKVASHVRGCACHLRPCIRFCCPQYQKMQKSKCYGDMSEDEL
NKHDPFVNVTLSDGSVVRRHFKEDLIVQSDLAKPGCPRMYFLNHELPGNEFTLFENGSLLRHWDKVELSKPGVLRPASFI*
(SEQ ID NO: 54)

Name: mth-like 3
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384666

AATTAATGCATTTTTGAAAATTATGGGCGTGACTGATTTGGTCGGTTAGTGGACGTGTCAATATTATGAAGCAAACGTGGACACTACAAACTGAAATATATATATATATATC
ATTTCATTGCCCTTAAGTCTGAAAAACGGCTTTAAAATTGAGAAGTTACTTTAAGAGATATTCGAGGTATTCGGTGACAGCAGCGCACATTATCGACTGGCCAATGGAATA
TTAGATTTGTAGAATAGGGTAATATGTAAACGTAAACCTTTCGATTGTATAAAGCATTCAATTCCTAATATGGATATGTCCGTTCACCAGCCTGTATTAACTTTGCGTGTTA
TTACTTAAATTTATCTTTCTATGTACATGCCAGAAACTGATAGAATGGCACGGTTACCGCCGATAGTCGCTGAGTATGGGGTATCTTATAGTCGAGCACATTCTCGTTCTCC
CTTGACATCGCATATATCGGTTGAGTTGTTCGCATTTTGGGATCAGTCATCTTAAAACTCTGGTCGCATAAACATGTTTTTTGTGTTAGTCGTTCCATTTGACCCTTGCGTA
AGTATTTATATTCAAAGTTTTTACGATTTTGGATCGTTAAGCATAAGAAGGACCACTCGAGTACGATTTGACCGTTTTTATACAATTTCGTCCTCTTAGACGAATAAATATT
GTGCCTGCCTTGTTTCTAAAAATGAACTTGCTTCAGCTTCGTAGAGAATGCTTAGTCATAACAGATGGATTTGTTGAATTTATTTTTAACAGTTGATAAGTTATACTTTTAT
AAAAAAAAAAAAAACAACTTACTCAGAATGCTATAGCACATGACTCAATTTTTTGTTTTTCATAAGTATTTAGATCAGTTTTCAGTTTTGCGGAGGTTATTTTGCTACTTTTAT
TTCTAATTGTGAACACGTAACTTGGTTTAAAACCGTATATACTTACAGAATAGCATACATCATTTGAATTCAGTTGCCTATCATATTAATTTCCTTTGCAGAAAATGCGGAT
ATTGTTAATTGCAGTTCTTTTTCTGTTGATGCCAAAATCAAATGCCGAAATTCCCGGTTGCGACTTCTTCGACACCGTAGATATTTCAAAAGCGCCAAGATTCTCGAACGGA
TCGTACCTCTACGAAGGCTTGCTGATCCCCGCCCATTTGACAGCTGAATATGACTACAAGCTCCTGGCCGACGATTCGAAGGAGAAGGTGGCGAGCCACGTACGAGGATGTG
CCTGCCACCTCAGGCCATGCATCCGGTTTTGTTGCCCCCAGTACCAAAAGATGCAAAAGAGCAAGTGCTATGGCGACATGTCGGAGGACGAGCTGAACAAGCACGATCCCTT
CGTGAACGTGACGCTCAGCGACGGGTCGGTGGTCAGGAGACACTTCAAGGAGGATCTGATCGTGCAGTCGGATCTGGCCAAGCCCGGATGTCCCCGGATGTACTTTTTGAAT
CACGAACTGCCGGGCAATGAATTCACTCTGTTTGAGGTTACTTTCCTGTGTTTCTGTCACTTCTGAAGATCGTAACCTTTTTTGAATCTCTGTAGAACGGCTCTCTTCTGCG
TCACTGGGACAAAGTGGAACTGAGCAAGCGGGAGTACTGCGTCCAGCATCTTTCATTTAAGGATGATAGTATCCGAATTGCACCCCACTTCTGTCCACTTTCGTCTGAGCAC
TCCAGAACGTGGAAGACCGTTGGTAAGCTGAATTCCCTGCTTGACTTTAAGGTTTTTAGTTAATGTCCTTTCTCACAGCGATAGTGATATCCTTGATATGCATAATCCTAAC
GATCAGCGTGTACCTCTATGTCGAGAAGCTTCGCAACCTGCACGGAAAGTGTTTCATCTGCTACTTGGCCTCTTTGTTCTTGGGCTATTTCTTCCTGGTCCTCAACGTGTGG
AAATACTCCTCCGGCTTCTGCGTTACAGCAGGTTTGCTTGTGGCGATTACAAGATAAAGGACGACTAATTTCACATTTCGATATCATGATCAGGGTTCCTGGGCTACTTCTC
CGTCATGGCCGCGTTCTTTTGGCTTTCTGTGATTGGAATCCACTTGCGGATTAAGTTTAGTCTAGCCTCTAACTGTTTGCACCGATTGCTGCCAGAAAATCCGTTCCGTGCA
TACAACTTATACGCCTGGGGTATACCGCTAATCATGACTGCAATCACCTATACAGCTGACCAGGTGGTTAAGAACGAGAAGTTGAGACCCCGCGTGGGCGTCGGCAAAAATT
GTTGGATCTACAGTAAGTTACTATTCCTATAATTACAATTACGGTGTAACTCGTTTGTTAACCAACAGCTGGCGATATGACTGTCATGATCTACTTCTATGGCCCGATGCT
GTTGCTAATCGCTTTCAACATAATAATGTTTGTCCTTTCGGCGATTTACATATATAACATAAAGAAAAATGTGAAAGGCCTTGTCCACAAGCAACAAACAAACCAACAGATA
AATGACCAACAAATGTAATTATATTATATTATAAGATAATAACTATAATGTCAAATGGCACTTTTTGTAATCTTTTATAGGTTTGCTATATTCCTGCGACTTTTCATCTTAA
TGGGTTTGTCGTGGAGCTTTGAGATATTATCCTTTTTGTTAACCAAACAGCAAGCTTGGGCTAGGGCTTTAATGGTGGCTGACTACTTTAATTGGTCCCAGGGTACCATCAT
ATTCGTGCTTTTTATTTTAAAGCCTAGCATTCTAAAACTTATAATAGCAGGGTAAGTAAAGTATTACAAATAACCTTGAAATTTGACCGTCTCATACGTATGAACAATCTTT
TGATTTAGGGGACGTCAAAACCTCCCAGGAAGTCACCATAATTCGAGATCAAAAGCAGCTCGATATAATTCAACTCATACGGCTTGTGAAGGATCAATTGCGGATCCAAACG
CTTACTGCTAAAAAAATGGAATTCAATAAGATGCAACGTCTCTGTTATTGATCACATTATTTTAATTTTCTATGGAGCAAAAGTTTAAAAAAAATATTTGTTATCAAGACAA
TAAATAATGTCTTGGAACATATAAAGCAAAATATATTAAAGGGTACAAAAAACAGGAAAAGTATGGGGCAATACAGCAAAATATTTTTCGGATCATCAACTTTTTTGTAAGA
CAAAATGTAGAAATATTTCTGGCACCGTTGTTTCAAACCAGGGTTTGAGTCAACAATGCTGGCCCACCACTGGCAAGTACCAAACGATTACAGAAAACTTTTTCTAATCGAT
TAAGCATCGGGGGTCATCGATGTTTCTCCACCACTACAATTTCATGTTGTTTTTTAGTCGCGGCTGTTTGACGTTTGTCGCTTTCCTCCGTTGCGAATATAATATATTACG
TAGCTCATTTTTATACAAACGGAATTACGAGCGCAACGACGACAGCAACACTAGTAGCACTAATCGTAAGCGCAGGGGCCAAAAATTAAATTGCGTTTGCGGCCGCAAAGAT
TTGATGACGTCGCATACGCCGTCTTCTAGGGCGTAAAAAGCAAAGCAAAGCAAACAAACGCGAAAGCGAAACGTGTAAACGGCGTAGAAGCGATAAACGCGACTCAAATACG
CAGCAGATAAAATACAATACGCGAGAAGAGAAAAGTCACGGGAAATATTGTTCATA
(SEQ ID NO: 55)

Exon: 1001..1492
Exon: 1552..1702
Exon: 1759..1935
Exon: 1998..2252
Exon: 2627..2640
Start ATG: 1001

Transcript No. : CT20351
ATGCGGATATTGTTAATTGCAGTTCTTTTTCTGTTGATGCCAAAATCAAATGCCGAAATTCCCGGTTGCGACTTCTTCGACACCGTAGATATTTCAAAAGCGCCAAGATTCT
CGAACGGATCGTACCTCTACGAAGGCTTGCTGATCCCCGCCCATTTGACAGCTGAATATGACTACAAGCTCCTGGCCGACGATTCGAAGGAGAAGGTGGCGAGCCACGTACG
AGGATGTGCCTGCCACCTCAGGCCATGCATCCGGTTTTGTTGCCCCCAGTACCAAAAGATGCAAAAGAGCAAGTGCTATGGCGACATGTCGGAGGACGAGCTGAACAAGCAC
GATCCCTTCGTGAACGTGACGCTCAGCGACGGGTCGGTGGTCAGGAGACACTTCAAGGAGGATCTGATCGTGCAGTCGGATCTGGCCAAGCCCGGATGTCCCCGGATGTACT
TTTTGAATCACGAACTGCCGGGCAATGAATTCACTCTGTTTGAGAACGGCTCTCTTCTGCGTCACTGGGACAAAGTGGAACTGAGCAAGCGGGAGTACTGCGTCCAGCATCT
TTCATTTAAGGATGATAGTATCCGAATTGCACCCCACTTCTGTCCACTTTCGTCTGAGCACTCCAGAACGTGGAAGACCGTTGCGATAGTGATATCCTTGATATGCATAATC
CTAACGATCAGCGTGTACCTCTATGTCGAGAAGCTTCGCAACCTGCACGGAAAGTGTTTCATCTGCTACTTGGCCTCTTTGTTCTTGGGCTATTTCTTCCTGGTCCTCAACG
TGTGGAAATACTCCTCCGGCTTCTGCGTTACAGCAGGGTTCCTGGGCTACTTCTCCGTCATGGCCGCGTTCTTTTGGCTTTCTGTGATTGGAATCCACTTGCGGATTAAGTT
TAGTCTAGCCTCTAACTGTTTGCACCGATTGCTGCCAGAAAATCCGTTCCGTGCATACAACTTATACGCCTGGGGTATACCGCTAATCATGACTGCAATCACCTATACAGCT
GACCAGGTGGTTAAGAACGAGAAGTTGAGACCCCGCGTGGGCGTCGGCAAAAATTGTTGGATCTACACAAGCTTGGGCTAG
(SEQ ID NO: 56)

Start ATG: 1

MRILLIAVLFLLMPKSNAEIPGCDFFDTVDISKAPRFSNGSYLYEGLLIPAHLTAEYDYKLLADDSKEKVASHVRGCACHLRPCIRFCCPQYQKMQKSKCYGDMSEDELNKH
DPFVNVTLSDGSVVRRHFKEDLIVQSDLAKPGCPRMYFLNHELPGNEFTLFENGSLLRHWDKVELSKREYCVQHLSFKDDSIRIAPHFCPLSSEHSRTWKTVAIVISLICII
LTISVYLYVEKLRNLHGKCFICYLASLFLGYFFLVLNVWKYSSGFCVTAGFLGYFSVMAAFFWLSVIGIHLRIKFSLASNCLHRLLPENPFRAYNLYAWGIPLIMTAITYTA
DQVVKNEKLRPRVGVGKNCWIYTSLG*
(SEQ ID NO: 57)

Name: mth-like 4
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384544
GATTGCCGTGGTCGTGCGCCGCGGGAGAAATGCCCTTTTCGTGTACGCAATTACTAGGCAAAAAGGTTTGGACGGCCCCTTGCTTTATTAGGCCATAATTTACGCTGTCGAGC
GGCCATTAGAGCGGCATCCGCTGGCCAAGTTGGTCTATAAAAGCCCGGCGGCTGCAAGTGGCGTTTTCCATTCCACGTGAAAGGACACTCGCAGTCGGGCCGACGAGATGCT
GCGCCACCTGCTCCGCCACGAGAACAACAAGGTCTTCGTCCTGATCCTGCTCTACTGCGTCCTGGTCAGCATTCTGAAACTCTGCACGGCACAGCCGGATAGCTCTGTGGCC
GCCACGGATAATGGTATTTGATAAGGCATTTGATAATATCATATTATTCGCTCTAATGGATTCCATTTACCATCCAGATATTACGCATCTTGGCGACGATTGTCAGGTGACG
CCCGTCATCCATGTGCTCCAGTATCCTGGATGTGTGCCCAAGCCGATTCCCTCGTTCGCCTGCGTGGGTCGCTGTGCCAGTTATATCCAGGTGGGTGTTGATATACTTGAAA
TATAAATTATAATTTGATATAATCTGATGACTTTAGGTTTCGGGCAGTAAGATCTGGCAAATGGAGCGTTCCTGCATGTGCTGCCAGGAGTCTGGTGAGCGGGAGGCAGCCG
TCTCGCTATTCTGTCCCAAAGTGAAGCCCGGCGAGCGTAAATTCAAGAAGGTCCTGACCAAGGCGCCATTGGAGTGCATGTGTCGGCCATGCACTTCCATTGAGGAGTCTGG
CATCATACCACAGGAAATTGCCGGCTATTCGGACGAGGGTCCACTCAACAATCACTTCCGGCGCATTGCTCTGCAATAGATTCCCCCATCAGTTTAGCACTCATACCCATGC

CCATTTGCTCATTAAAATAAATTAGAGTTGCCGTTTCTGCTTGTACACAGTGAAAAATGTGCCTCTGAAAGCTCATGCCAGTGAAAGGGAAACATCAGCAGATATATGATTA
TTTTAAATTGTAATTTATTTCCCAACTATCTCATGTCATTGCTACTAGTTCACTAAATGGTGTGTGTCCTGCTTCACAGGGTATGTATGTAGATGCGTAACTCCAAAGAAAA
CAAAGATAATCTTTACACTGCAACTGGATCTAACAATAGGAGTTTAACTACGGCATAATGCATATGACTCGCGCCGCCTACAAGTACTCGACGATATCGCATTTGATGGATA
TATTCGGACAGCTCTGGCCCACGGCACTGCGGCCCCGATAGTTGGGACTCTTCTCGGTGAGGAGTTTCCGGAAGATGGGCAAAATTGCCGTCTCGCCCTCGAGCAGTTCCGT
CTCGCTGCTGGCGCCGTAGATGCTGCTATCCTGCTTGCCGCGCTGCCTGCTCCGCTGCTTATCGGGCGCCACATTGGGCGAACTGTGGTGGTGGGCGGCGTGCTGGATGTTA
GAGCTCCGGCTGACCTTACGCGGCTGGACGCACGGCTGACTGTTGGGCGGCGGATCTCGTTCCCGTTCGCTGTCCGTCTCGCTCTGGCAAACGGTGCCCTTTTGCGGTTTTG
GCCGATGTTTGTTGCGGTTCTTCTCGGAGACGGACCGCTGGATGGGCGAGGTGGTGTTCGGGTCATAGGTCATCCCGTGGCCGCCATGATGTTCCGCACAGCGCTGCGGACT
GGCGTAGTTCCGCGGCGACTGGCGCTGCTTGGCGTGCGACAGCGTGTGCGTGTCCCAGAGGTCACTCTGTGAAACCACGAAATTGGACTTCAGAGCGGAGCCCATGCTGGTG
CGGCGATGGTCGGAGGCAGTGGGAGCTATTTTCGATTAGAAAATCGATAGTATTATAGTATTATAAGAGACGCGCAGAAAACTTACTCTTAATCACAGCCGGAGTACTGGGC
TTTGAGTGTGGACACATGTTGCTCAACTCCGATTGCGATGTGGAGATGTTGGTGGAGGAGTTCCGGTGCGAGCTGCTGTACTTTGGCCGGGCGGTGGGCGTCTGGCTGCCAC
TGTCGTAGTGCCTGGAGGCGGTCGATTGACGCCTCTCCCGCATGCTGGAGCAGTCTCCTCCGCTCCGCTTGCCCGCCGTACTGGTGATGCTGCCTATATTATCATCGGTCTG
AGCGGACACCGAGTTTCTCTTCTCCAGATGGCGATGATGATGCTGCTGCTGCTGCTGTTGCTGTTGCATCTGCTGGTGTTGCTGCTGCTGCAGGTGCTGCTGCTGCATCTGC
TGATGTTGCTGTAAGTGCTGCTGCAGTTGCTGCTGGATGGGCTGCATCATAGTTGGAGCAGGGGCATGCGAGTGGTGCTCTACAATCGATTGCTTTTTCTTTTTGCTGTTGG
AGAATATCCATAGATTAGAGTGTATTTAAGACCAAAATTTAAGAACTTACGGAGGAAGCGATGGCGCCATCACTGTAGACTGCACCGACACAAAGGAGTTATCCATCTCCAC
ATGAGTCGAGGTCACACTATCCACGGACGTCATCTCGGTAACTTCTATGACAAACGGGTTGCAATGCATGAACATATCTACGAAATCCCCCTCCTTAGCCACTCACCTGTGG
TGCTGGGCATCTCTCGACGCACCGGTGGAGCTGAGAGCCTAACAATGTCATCGTTCAGTATGGGCTCCAGTTCGCTACCGACGTTGGAGCCACCTGTGCAGGTCACCCCATC
GATCCATTTGCGTGCCTCGGGTCCCAGCTTGCGGATTAAGGCCTGCAGCTCGTTCTCCTTCTCCATCAGCGCCTTTCGGAATCGGCAGTTTGTGTTCTTTACATCTCGCAAT
CGTTGCTCCAGTTCGCACACCGAGGAATCCCGGCGTCCGTTTTTGGACATGGGTCTCAACGTGGCCCTAACGCGTTTGTCCACCACGCCCTGGGGATTCCGCTTCAGCTCCA
CCAACTGCAAAACGATTCGCATGAATATTAACCAAAGATTATACTACGCGGTCAACATCTAGGAAGGATCAACTCTACGGACAGTAATATAGGATAGTTAAAGATATTCCAA
CAGCAAGGAAGTGGGATTTGCCTACAACCTTGCTGCTGCTCTTAGTGGGTTCTATATAGTTCAAGTTTATAGTTAGTCTGTAGTGTTATATAGCATTTTAAGATCTGTTTTC
GAGTACATGTTGTTTCTGTATCAAGGTGTATCAACGCTAGACGAAAGATCAGCTCAGCTGGCGGTAAAGTACGATCAGTTAAGAAGCTCTAACCACAACTGATCGGCCTATT
ACCGCCGACAAACAACATTCATATTACACGATTATTATATGTAACATATAACAGCTACAAAAACTACAAAACTTCCTTACTACCAGTCTTACTTTCGGTACGAACACCAAAC
ACAAAGTGGCTGTCGTACAAAAAATGATAAAAAACGAGAGTAAGACAAAAACTAAATCCTTTCGATCCGATAGCACCAGGGATATAGCCGCTCCGGCCAGACAAGTGATGAA
CACGTTATAAACGGAGAAACCAATATGCTTGGAATCGTTCAGAGCCGGTATAGAAACATGTCGAGTTTCCCAGGCCAAAAAGGCGCCAAAAACCTATTGAAAAAAATTATTG
CTTTATTATAATGATATCTTACAAATTAATGTTTATTATATACTCACCAACAACAGTCCCTTGTAGGCATAAATAATGCTAACGAATATGGTCATGTGCTCAGACTGGCAGT
ACTCGTTTTCGGGGATCACCAAGACATCATCAATATTCTCGTGATGCTGTAGTTAATAATAATCATTTTAGATTTTTCACTTCAAATATGTTCTATCAATGCACTCACCAAG
GGTTCCAACTGTTTAGTTTCGCGGTAAAAGGGATCGGCAATCTGCCAGGTGGTTATAATGGCTATATCAATGGCCAAAAGCACGCCCACAACCATAAACAATTGATAGTCCT
TGATCACCTTCTTATTGAGCTTCAGATCGGTGAATATCGAATGCACCCGCCACGTCTTCGAGAACATGGCTCCAAAACTGAGACTGAATCCAGCCATCAAGATCCAGGCTCG
AGCTGTGCAGATATAGGGAAAAGCTGCCACACTACTTAATGTGGTATCGAGACCCAGGAAAATGATGCTCAAATAGGTCATCATACAGCCCACAATGATCAGATTGTTCAAA
TGGGGACTGGACATCTTGATGTATCTATAAAAACATCAATATTTTATTAAACGAAATTGTTTAATATAGCATTGTTAAGCTCACCTTTGATTGCGATACTTAATGTTAAAGG
CCAGAAAAACTGTGGCAATAATCACTCCAATGACCGAAGCACTAGCCGATACAATATATATGGTTGGATTGACCTGACTGTGCTCGATGTAGATCAAAGTGCGATCCTTGGG
AGGAGTTTTCCCCACCCATTTGACTGGTTTTCCCAAGCTTAAATCCAAGTGTGACTTCTGTGAGTGGTATTCCCCGATCTTTTCCATTTGTCCCAGCTGAAACTGATTGATC
AGGATGTTGGCCTTGCGCTCGTTGTTGTAGAAACGCACGGGTCCCTTGAGAATGTAAAATATGTTTTAACTGCAAAAACTGGATTGTTTATCTATCTACCGTCACACCCTCG
AAGGATGTATTACGTAGAGCCTCAAGGAAGACACTCTCCCAGTCCTTCACGCGATAATCAAAATGTGTTAGCAGATCCTCTCGCTTTTCGGCCACATACTGAATGGCCAGGG
CAGCTGCCCAGATACCATCGTAGGTATAGCCATGAAAGCGGGAATATTCAGTGCCTCGCAGTCTGTCGTACTCCACAAGATACTCATCAGCAGTCTAGTAAAATATGAAATT
TATTACAAGAATTTTAATAGACTTTACTGGTACTAACAATGCCAGCCACTGTGATGTCACCACTGGTGGACAAGGGCAAAAGATCCACTAGAATGGCACCTTCCAAGGCTGT
AGCGATCTCCTCCACACTGCACTCGCTGTCCTGCGTGACATTCCACCAATCCGTGGAATAGGTAGCCATGATCAGCCATTGATAGGCTCTGCCATACATATCCAATCTGGAA
GGGAGTTTATTGCAATAATTTTAAATTGAGACTTAAGTTAGTATATCCAACTCACTTATAAGCCTCACAGAATGCCTTGCGTGCAAAGTGCTCGTTAAAGTTGCCCAGAATG
ATCCTCACGTCCTTCTCGCGCAGTTTCTTCAATGATTCAGCCACATCGTTGACGAAGCTCTGCGTTTCCACCACCTCGACCTCCATGGCATCCAGGTCAGCCACCATGTGAT
TGTGGGGCAGCGAATAGCGTGGCTCATTCTGGTAGACAGTGCCCACTCTGGTCCAATTGAACTCCTTCAGCAAGGCCAGTCGCGGCGCATTAAAGGCATTCTCCGAGGGTAC
CACGCGAAAGAAATTCGGAAACGCATCCTTGGTGAACATGGGATGGGTGTCCGCGTAGCTGAGCTGGGTGAGGTGCCAGTGCTTGCTGGCCTTGGCTATGGGATCGGTCACA
TGGGTGCACGCAGCGCCGAAGAGCATCACTTTATTGGGACCCGAATGCATCATATCGAAGAAGGACTTTACGCCCACAGCAGCATTGCACTGGAAAGGAAACCAAAATTACA
TTCCAGAATATATGGAAATATGCATAGGTAAAGGTGTGATATTACAATATGAACCAATGAACATTAAAAAAAATACCTAATTGAACTTTAAAAACCACAAGTCCATCATCAG
AGATTCACTTAATGAAATCTTCAAGCTAAACCCTAACTCACCTGAGTGTCGTTCCACCACATGTGCAGCCTGTAGTTGGCCAGTATCTTTCCATGCTCATTAACGTGACCCA
AGGCGAGCTTTACACTGGGCATAACGCCCCGACCGGTGTAGGAGTTTTCCACGCCATCCCCGTACGGGAAGAATCCCGCTATGTAGACGTCCGATCTCTTGGCTGTCCTGCC
ACAGGCGGTGCTCCAAAGGAGCAGGAACAGCAGGCTGGCGAATGGAAACCAACTTGGCCGGAACATGTTGACCGCGTGGGCTGTAAAACGATTCTAAATGTGAATAGCGCAT
ATCGAGGACGTGAACTTGGAGCCAAGTTGTGGGTGGTGAATGGCAGTGGTTGGTTTCGGGTGAGCCAAGGCTGCCGAGGCAAGGCTGCGCCCAGTGGGTGTGTATTATTTTC
TACGTTCTACCTGTGTTTGCTGGCACTCGCATGCAATTTTACAGGACAGGTGGAGCAAGGCTAAGGCCAGTCCAGGCCAGCAAAATTGCTTCAAATTGCTGCAGCAATGCTG
GCACGTTTAGTAATTATAAAGCAATTTAATGATCTAAAGGATTCGTGCATTATTTAGTATGCTCGTTAACGCGAAGGAGTATCTTCCAGGTGTTGCATCCTCTTAAATCTTA
TGTGTACATGTTTTTAGGAACCTACCAGCTTGCTTAGTCCCAGAACTGGCTTTTATTTCACTACCGATGCCATTTGAAACACTTATTTTACATCAAATCTCAGCCAACGAAG
AACTTTTGCAGTGCGAGGGTCCGAAAAATTATGTATCCCCCTGACTTGACATTTGTCTGACAGGAGGAATTCGAGGCAGGAAGTGCGCATGTCCTTAAACCGAAAGCTGTCA
CTGCTCCGCCTTGGCCTAGGCTTTCCAACCCCCGTAATTACCATAAACCCACTGCGCTTGTAAACATCCTGCAACAGGTGGTCCCATCAGCTGGTACTCCGGCAATTGTTGT
GCCCGTTGCCATGGACTACTGGGGTCGAAAAACGAAGTAGATAACTGCTCGGAAAATTACCAAATTGAAAATTGTGAGTAACCATAATGGAGGACTATGCGCCCAGCCGCTA
TAAAATGCTTGAGAAAATCGGCGAAGGCGTCCATGGCTGCGTGTTTAAGGCCATCGATCTGCAGCGCAACAAGGAGGTGGCCATCAAGAAGGTGGCCCTTAAGAACAAGTTC
GGAAACATAGCCCTAAATACTTTGAGGGAAATCAAGACCCTGCAGCTTTGCAAGTCTGAATATGTAGGTATACCATATATCTATTTTCCATTTGTTAATTATGTAACTAATT
GTAGATTCTGGACATCATTGATATATATCCTGACCTAACTGGTCTCTCCTTGGTGCTGGAATACCAGCCGGACACTCTGTACAACCGCCTCAAGAGCGAGGTTAATCCCCTG
AGCCGCCAGCAGGTGCGTAAGTTTGCACACCAAATGTTCAAGGGGATTGCATATCTGCATGAAGCGGGTCTCATGCACCGGGATATTAAACCCGCCAATCTGCTTATCAGCG
ATACAGATATGCTTAAGATTGCCGATTTCGGACTGGCCCGTCTGTACTTCCCTGAAGACGAGTCGCGCCTTTATTCGCCTCAGGTCTCAACCCGGTGGTATCGAGCGCCAGA
GATTCTATTTGGCAGCCAGAAGTATGGAACCGGAGTGGATATGTGGGCGGCAGGTTAGCCATGACTCTTTGATGGGTTTACATATATCCTTATGGCTGACATTATTGCTAT
(SEQ ID NO: 58)

Exon: 6503..6410
Exon: 6016..5754
Exon: 5577..5096
Exon: 5034..4854
Exon: 4798..4580
Exon: 4524..4229
Exon: 4168..3805
Exon: 3742..3632
Exon: 3565..3341
Exon: 2916..2571
Exon: 2508..2403
Exon: 2345..1879
Exon: 1817..1001
Start ATG: 6002 (Reverse strand: CAT)

FIGURE SHEET 28

Transcript No. : CT20836
AAAAGTTCTTCGTTGGCTGAGATTTGATGTAAAATAAGTGTTTCAAATGGCATCGGTAGTGAAATAAAAGCCAGTTCTGGGACTAAGCAAGCTGCCCACGCGGTCAACATGT
TCCGGCCAAGTTGGTTTCCATTCGCCAGCCTGCTGTTCCTGCTCCTTTGGAGCACCGCCTGTGGCAGGACAGCCAAGAGATCGGACGTCTACATAGCGGGATTCTTCCCGTA
CGGGGATGGCGTGGAAAACTCCTACACCGGTCGGGGCGTTATGCCCAGTGTAAAGCTCGCCTTGGGTCACGTTAATGAGCATGGAAAGATACTGGCCAACTACAGGCTGCAC
ATGTGGTGGAACGACACTCAGTGCAATGCTGCTGTGGGCGTAAAGTCCTTCTTCGATATGATGCATTCGGGTCCCAATAAAGTGATGCTCTTCGGCGCTGCGTGCACCCATG
TGACCGATCCCATAGCCAAGGCCAGCAAGCACTGGCACCTCACCCAGCTCAGCTACGCGGACACCCATCCCATGTTCACCAAGGATGCGTTTCCGAATTTCTTTCGCGTGGT
ACCCTCGGAGAATGCCTTTAATGCGCCGCGACTGGCCTTGCTGAAGGAGTTCAATTGGACCAGAGTGGGCACTGTCTACCAGAATGAGCCACGCTATTCGCTGCCCCACAAT
CACATGGTGGCTGACCTGGATGCCATGGAGGTCGAGGTGGTGGAAACGCAGAGCTTCGTCAACGATGTGGCTGAATCATTGAAGAAACTGCGCGAGAAGGACGTGAGGATCA
TTCTGGGCAACTTTAACGAGCACTTTGCACGCAAGGCATTCTGTGAGGCTTATAAATTGGATATGTATGGCAGAGCCTATCAATGGCTGATCATGGCTACCTATTCCACGGA
TTGGTGGAATGTCACGCAGGACAGCGAGTGCAGTGTGGAGGAGATCGCTACAGCCTTGGAAGGTGCCATTCTAGTGGATCTTTTGCCCTTGTCCACCAGTGGTGACATCACA
GTGGCTGGCATTACTGCTGATGAGTATCTTGTGGAGTACGACAGACTGCGAGGCACTGAATATTCCCGCTTTCATGGCTATACCTACGATGGTATCTGGGCAGCTGCCCTGG
CCATTCAGTATGTGGCCGAAAAGCGAGAGGATCTGCTAACACATTTTGATTATCGCGTGAAGGACTGGGAGAGTGTCTTCCTTGAGGCTCTACGTAATACATCCTTCGAGGG
TGTGACGGGACCCGTGCGTTTCTACAACAACGAGCGCAAGGCCAACATCCTGATCAATCAGTTTCAGCTGGGACAAATGGAAAAGATCGGGGAATACCACTCACAGAAGTCA
CACTTGGATTTAAGCTTGGGAAAACCAGTCAAATGGGTGGGGAAAACTCCTCCCAAGGATCGCACTTTGATCTACATCGAGCACAGTCAGGTCAATCCAACCATATATATTG
TATCGGCTAGTGCTTCGGTCATTGGAGTGATTATTGCCACAGTTTTTCTGGCCTTTAACATTAAGTATCGCAATCAAAGATACATCAAGATGTCCAGTCCCCATTTGAACAA
TCTGATCATTGTGGGCTGTATGATGACCTATTTGAGCATCATTTTCCTGGGTCTCGATACCACATTAAGTAGTGTGGCAGCTTTTCCCTATATCTGCACAGCTCGAGCCTGG
ATCTTGATGGCTGGATTCAGTCTCAGTTTTGGAGCCATGTTCTCGAAGACGTGGCGGGTGCATTCGATATTCACCGATCTGAAGCTCAATAAGAAGGTGATCAAGGACTATC
AATTGTTTATGGTTGTGGGCGTGCTTTTGGCCATTGATATAGCCATTATAACCACCTGGCAGATTGCCGATCCCTTTTACCGCGAAACTAAACAGTTGGAACCCTTGCATCA
CGAGAATATTGATGATGTCTTGGTGATCCCCGAAAACGAGTACTGCCAGTCTGAGCACATGACCATATTCGTTAGCATTATTTATGCCTACAAGGGACTGTTGTTGGTTTTT
GGCGCCTTTTTGGCCTGGGAAACTCGACATGTTTCTATACCGGCTCTGAACGATTCCAAGCATATTGGTTTCTCCGTTTATAACGTGTTCATCACTTGTCTGGCCGGAGCGG
CTATATCCCTGGTGCTATCGGATCGAAAGGATTTAGTTTTTGTCTTACTCTCGTTTTTTATCATTTTTTGTACGACAGCCACTTTGTGTTTGGTGTTCGTACCGAAATTGGT
GGAGCTGAAGCGGAATCCCCAGGGCGTGGTGGACAAACGCGTTAGGGCCACGTTGAGACCCATGTCCAAAAACGGACGCCGGGATTCCTCGGTGTGCGAACTGGAGCAACGA
TTGCGAGATGTAAAGAACACAAACTGCCGATTCCGAAAGGCGCTGATGGAGAAGGAGAACGAGCTGCAGGCCTTAATCCGCAAGCTGGGACCCGAGGCACGCAAATGGATCG
ATGGGGTGACCTGCACAGGTGGCTCCAACGTCGGTAGCGAACTGGAGCCCATACTGAACGATGACATTGTTAGGCTCTCAGCTCCACCGGTGCGTCGAGAGATGCCCAGCAC
CACAGAAGTTACCGAGATGACGTCCGTGGATAGTGTGACCTCGACTCATGTGGAGATGGATAACTCCTTTGTGTCGGTGCAGTCTACAGTGATGGCGCCATCGCTTCCTCCC
AAAAAGAAAAAGCAATCGATTGTAGAGCACCACTCGCATGCCCCTGCTCCAACTATGATGCAGCCCATCCAGCAGCAACTGCAGCAGCACTTACAGCAACATCAGCAGATGC
AGCAGCAGCACCTGCAGCAGCAGCAACACCAGCAGATGCAACAGCAACAGCAGCAGCAGCAGCATCATCATCGCCATCTGGAGAAGAGAAACTCGGTGTCCGCTCAGACCGA
TGATAATATAGGCAGCATCACCAGTACGGCGGGCAAGCGGAGCGGAGGAGACTGCTCCAGCATGCGGGAGAGGCGTCAATCGACCGCCTCCAGGCACTACGACAGTGGCAGC
CAGACGCCCACCGCCCGGCCAAAGTACAGCAGCTCGCACCGGAACTCCTCCACCAACATCTCCACATCGCAATCGGAGTTGAGCAACATGTGTCCACACTCAAAGCCCAGTA
CTCCGGCTGTGATTAAGACTCCCACTGCCTCCGACCATCGCCGCACCAGCATGGGCTCCGCTCTGAAGTCCAATTTCGTGGTTTCACAGAGTGACCTCTGGGACACGCACAC
GCTGTCGCACGCCAAGCAGCGCCAGTCGCCGCGGAACTACGCCAGTCCGCAGCGCTGTGCGGAACATCATGGCGGCCACGGGATGACCTATGACCCGAACACCACCTCGCCC
ATCCAGCGGTCCGTCTCCGAGAAGAACCGCAACAAACATCGGCCAAAACCGCAAAAGGGCACCGTTTGCCAGAGCGAGACGGACAGCGAACGGGAACGAGATCCGCCGCCCA
ACAGTCAGCCGTGCGTCCAGCCGCGTAAGGTCAGCCGGAGCTCTAACATCCAGCACGCCGCCCACCACCACAGTTCGCCCAATGTGGCGCCCGATAAGCAGCGGAGCAGGCA
GCGCGGCAAGCAGGATAGCAGCATCTACGGCGCCAGCAGCGAGACGGAACTGCTCGAGGGCGAGACGGCAATTTTGCCCATCTTCCGGAAACTCCTCACCGAGAAGAGTCCC
AACTATCGGGGCCGCAGTGCCGTGGGCCAGAGCTGTCCGAATATATCCATCAAATGCGATATCGTCGAGTACTTGTAGGCGGCGCGAGTCATATGCATTATGCCGTAGTTAA
ACTCCTATTGTTAGATCCAGTTGCAGTGTAAAGATTATCTTTGTTTTCTTTGGAGTTACGCATCTACATACATACCCTGTGAAGCAGGACACACACCATTTAGTGAACTAGT
AGCAATGACATGAGATAGTTGGGAAATAAATTACAATTTAAAATAATCATA
(SEQ ID NO: 59)

Start ATG: 109 (Reverse strand: CAT)

MFRPSWFPFASLLFLLLWSTACGRTAKRSDVYIAGFFPYGDGVENSYTGRGVMPSVKLALGHVNEHGKILANYRLHMWWNDTQCNAAVGVKSFFDMMHSGPNKVMLFGAACT
HVTDPIAKASKHWHLTQLSYADTHPMFTKDAFPNFFRVVPSENAFNAPRLALLKEFNWTRVGTVYQNEPRYSLPHNHMVADLDAMEVEVVETQSFVNDVAESLKKLREKDVR
IILGNFNEHFARKAFCEAYKLDMYGRAYQWLIMATYSTDWWNVTQDSECSVEEIATALEGAILVDLLPLSTSGDITVAGITADEYLVEYDRLRGTEYSRFHGYTYDGIWAAA
LAIQYVAEKREDLLTHFDYRVKDWESVFLEALRNTSFEGVTGPVRFYNNERKANILINQFQLGQMEKIGEYHSQKSHLDLSLGKPVKWVGKTPPKDRTLIYIEHSQVNPTIY
IVSASASVIGVIIATVFLAFNIKYRNQRYIKMSSPHLNNLIIVGCMMTYLSIIFLGLDTTLSSVAAFPYICTARAWILMAGFSLSFGAMFSKTWRVHSIFTDLKLNKKVIKD
YQLFMVVGVLLAIDIAIITTWQIADPFYRETKQLEPLHHENIDDVLVIPENEYCQSEHMTIFVSIIYAYKGLLLVFGAFLAWETRHVSIPALNDSKHIGFSVYNVFITCLAG
AAISLVLSDRKDLVFVLLSFFIIFCTTATLCLVFVPKLVELKRNPQGVVDKRVRATLRPMSKNGRRDSSVCELEQRLRDVKNTNCRFRKALMEKENELQALIRKLGPEARKW
IDGVTCTGGSNVGSELEPILNDDIVRLSAPPVRREMPSTTEVTEMTSVDSVTSTHVEMDNSFVSVQSTVMAPSLPPKKKKQSIVEHHSHAPAPTMMQPIQQQLQQHLQQHQQ
MQQQHLQQQQHQQMQQQQQQQQHHHRHLEKRNSVSAQTDDNIGSITSTAGKRSGGDCSSMRERRQSTASRHYDSGSQTPTARPKYSSSHRNSSTNISTSQSELSNMCPHSKP
STPAVIKTPTASDHRRTSMGSALKSNFVVSQSDLWDTHTLSHAKQRQSPRNYASPQRCAEHHGGHGMTYDPNTTSPIQRSVSEKNRNKHRPKPQKGTVCQSETDSERERDPP
PNSQPCVQPRKVSRSSNIQHAAHHHSSPNVAPDKQRSRQRGKQDSSIYGASSETELLEGETAILPIFRKLLTEKSPNYRGRSAVGQSCPNISIKCDIVEYL*
(SEQ ID NO: 60)

Name: GABA-B receptor like
Classification: G_protein_linked_receptor
Gene Symbol: BcDNA:GH07312
FlyBase ID: FBgn0027575

Celera Sequence No. : 142000013383823
GTTCCTTAAACGTGATCAAAAATATAACTAGAGCTACTACTAAGCAGCGAGCACAGCGTGCGATATATCAACAAGTCCAGCTGGCAACATGTTGCCGCGCCTGTGGTGAGTT
GTGCCTCTATGTCCTTTGGCAGCATGTTGCGGTCGCTTGGCGCCGGCTGGTGGTCACAAAAAGCGCCGACGCTTGCCGTCAGTGCTTCGCGAAGATTGCGAGGCGCGACACA
CATCGCGGAACCCGAACTCCATACGGCTGTGCCGACACTCAATCCCGACCTAAGCCCAATTTTTTGTTGCGTGAAGTTGACGAAGTGTGTTGTACTGCTGCAAGTGCAAGTC
CCCGATTGCTAGTGCTATTCCGGGATCACAAACGAGCATCCTTCTTCGGCCTCACGATTGACGCCTTCTACCACTACCTGCGCCAGGCGCTGCCGCTGGCCAAGGAGGCTGC
GATCCACCTGAATGCCAGCAACGAGATCTCCGCTGTCGGTGACGGGGTCACCATCACAGGGACACCCGGTGATCTTCTCAACTACAGCGGGCTGGAATTGGATCTCGGACTG
GATTTGGATCTTAATCTGGACATGGACCTAGCCACTACTCCCAGCAGCAGCACCTTGGCGCCAGCTGTAACCGTCAGGACACCGGGGAACAGGAGCGTGGTGAGGGTCAGTG
CCGATGTGCCCATCTGGGTGGTGCCGTGCTATTCGGCCATCCTTCTGTGCGCCGTGGTCGGGAACCTGCTGGTGGTGCTAACGCTCGTCCAGAATCGTCGCATGCGCACCAT
CACGAACGTTTTTCTGCTCAACCTGGCCATCTCGGACATCCTGCTCGGCGTATTCTGCATGCCGGTGACCCTGGTCGGCACTCTGCTAAGGCATTTCATATTCGGCGAGCTG
CTCTGCAAGCTCATCCAGTTTGCTCAGGGTGAGTGGTGGAAGCCAGTAGGCCAGTTCATTTATGTCTTGTCCGTGTCCAAATTAAAAGGTTATCCCAAACTTCCAACTGTGA
ATTCCGCATTTATTAGTCCCTATCCCCGTGACCGAGCCCCATAATCGCTTGTCAAATGGCCACGAACCAGCACGACCTCGAGCGGATTCGCCAGCAGATCGTGTTGGCCAAC
ATTCAGGAGCTGATAAAGAAGATGACACGTCGCTGCTTCGACGTATGTATCGCTATGCCGGAAATGGAGTTGCGCTCCACGGAGCGCGACTGCCTGGCCAACTGTATGGATC
GATTCATGGACTCGGTTCAGGTGGTGTCGAGCCAGTACTTCCGTCGCCGGCGTCGCCATCAGCAAATTCGTTTGTCCCGTTCGACCGCCTCCTCCGCATCACCACCAGCATC
CGCATCCGCATCCATGCCCAAATCCGCAGCAGCAAATGAATCTGAATCCGCATCTAGGGCCTCCAATGACGAAAAGGTCAAATAATGCTTAAGTCTAGACAGATAGCCACGT
GTGGAGTTTATTCCAAGTCGAAAACGAAGGCTGTTCCGTCGTAGCCCCAGATCCCCAGTGTACTTGGGTGATGGCCACTTGTCCACATTGGTCATCATCCCAAAGCTTTCGA
GGATGATTTCATACTTCGAGACCTTTGGGAAGGGCAAGAAGGGGTCGTCTAGCTGGAAATGAAAGTGGATTTTCTGCAGGCGTGGCCATCGATGATTTCTTGCTATTGGTTC

TTAACAAAATCCTAGTAGAGTCTTAAGTAAATGACGATTGGCAGAAAATAAAGATGTCGTTCAGTGTACATTCACTATGCTTTGAGATTCTCTTTAAAGTCACCCTTTTGAA
ATCACTTAGCTAAGCTTAATATTGTATAATATACAAAGTAGTGTAGTGGGAGAAGATAAGTTTTAAATGACCATACTGTATATTTTAGATTGTGCGCTTGTGCACATCACT
TTTCCACTTAGGAAAACTGAGTCTGCCATCGATTTTACGCACAATTTCCGCGCTTATTTATGGCCTAAATCATATTTCGTTGGCAAGAACGCACTATATCTCGTCTCCCAAT
CTGCATCCGCCTGAGAATGGCCGTATTTTATGATGGCCGCCCGCATTTGCGTAACGACAGTCAGGGTTTCCCCGACCAGCGAGCACCGGGTTTTGGGTTCCTCAAGGGTTCA
AAAGATTAAATAGGCCCACGCACACGACACACACACGCACACACACGCACTCGCACACTGTAAAATATGCGGGGGAAGCTCTTTAAAACCATTTCAACTCAATGACTGTGCCC
AAAACTCATGTTTGGTCCTTGAAAAGTCCAGTCAGATGGGTAAACATTCTCACCAGGCATTAATTTGTGTCTCCAGCTCAGATAAGGGTTGGTGATGTGGGTGGCGATATTC
ATCAGCTGAGCCCGGTTCAGCCATCTGTGATATGGTTCGCCTTCGGCGTGATTCCATGTCCCGTCGGCGGCAGTGGAACAGAATGAAGCAGAAATGTGGCCCTGACACCTGT
CACAAATGCTAATCGGCATTTCGGTTTCAAGGAGCCGAGCGGCATTAATTTGCTTATTTAATCAAAGTGCATGCCATCGGACATTCTAATTATTTTCATAACGGTAAGGAGA
GGTCCTCAACCTGTTTCACCTCCTAGGACTCCTTGGAGGCGCATCCAAACGGATGATGAAGCTCAGTCCAGGCCAGTCATCCTCCATTCCCGCAAGATTAATCCAGGCTGAT
CCCCATTTCCAATGCAAATTGCTCGCCTGCCGGGAGCTCTGCTTCTGCTCAGCATTCTGGTCAATTTCCATATGTACGCAGTGTCGATCTGGGAGCCTACTCACCCTTCTCA
CCCAGCTCCTGTTACCCTGTTCCTTGCTCCTTTGGCGTTGTAGCACTCAAGTCGCAGCCTAAGCAGTCTGGAATTGCTTTTAACCGTGGCAATATTGCGAAAATCACTCTGC
TATGCAGCGTCACGTCGCGATTTTCTTGTAGTGTTCCGTTCCGAAGCCCCTTTCCCCATCAGCTCCATCAGCTCCGCTAGCTGCTCCACATCCTTTACGTCCCCTCGCCGA
TCCTTCTGACACCCTTCCTTGCATTTTCATGACTTGTTTCCTGTTTCCTGCGTTTCCGTTTCAGACGTGCTGGTGATGTTGTTTGTCTTGCTGCAAGTTTGTTTTCGGTACT
GCCGGAACCGCTGTTCTCTCCTGGCCTCCAACACCACCCCCCCCCTATCCACTGCCCCGGCATCCTGACGTATCCTGGCGTACCCTGGCATATCCATCTTGGCCGCCCCC
GCACTAACTTATTAACAACATAATAACAATGGCTGGCAAACAAGTTGAGCGCTCAGGCACACACATGCCAATTTGTTCAGCAAGCACTCGGCGCTGTCTCGAGCCCCAGAAA
GGCCCAGAAAAGGAGCTGGCATTCAGTTTCCAAGGATTCCGCGCCCCCATCCTCATCCGCATCCGCATCCTTGTTCGCATTCCCATTCCTTCCATCTGTTGCACGGGGCGT
CGACTCGTGGGTCGTAAATCTGAAAATGTCATCCCTGGCTGGGCTGAGGCACCAATTTATGGCCATCATCACGTCAAGGTGCAGCGATGTCCTGACACGTCCTTCCTGGCTC
CCTGTCCCACCTGGCTCCTTCCGGAGTGGGCGTGAACGCGCCCCTGGATTTATGATCGTAATTATGCATTTGTGGCCAGTGGCATTTGCATTGGACTCCTGCCAGTGGCAAC
CAATCGCCACCCACTGCCATTCGATTTCATCATCTTGCTTGCAAGAGTAATGTTGGCATTGTCCATTACTTATGTAAATCGAAACCGAAGGATAAGCTCCTGTAAGTCCATA
TTGGTTATAGTTTCTGCATATTCGATCTACCAAGAATTTCAACGAAATTCACAAGCGAGGCAGCAATAACAAGCTCAAAGTTAGTCGGATTATGGGAGCTACTAAAAATAC
ATAGTTGTTGACTTTTGGCTCAGAATTAGAGGTAATAAAACACAAACCGATGGTCTTTTTTATGGTCCTTTTCTTTTCGCAAGAAATTCCTCCTCGAAAACGAACATCTTA
GCAACTAATATTAGCAACTTTGTATGTTGCCCATACCAAAAAGAGTGTGAAAGGAATGTTTTCTGTTTGATGAACACTTTAAGGCTAGGACATAATCTCAGATAAACTGTTG
TCTAATTAACTTAATTTGATATCCAATTAAAGTAACTATGTGTGCTGGTAAGACAAAGAAAAGGTTTCTTTGTGGCGCTTGAAAAGGACATTTGTGTGCGGAAAAGGGTCAA
ATTGACATCCTTAGACCAAAGTTTATGGCTTTGGGCTTTAGAATTTTCTCTATTCCACTGCTCGACCTGGTTGCACTTTTTGTACCCCAGGCGGGACGGCCTTGAACCCTGA
CCTGTGGATGGGCCACCCATTGGGCGCGTCCAACTCGTCCAAGTCGCTGACAGCTCGGCAAATTAATGATTAATTGAATTTTCCTTGGCTCTGTTTTGGGTTTTTTTTTTT
TGTTGTCCATTCTATTATTTCTTGTTTTCTTGGCAACACAAATACAAATCAGCGCCACACAGACGTAATCATTAATTAGGCGCGCAGGGAGCTAAGTGTCTCAAGAAATGTG
GGCGCAGTGCAATTCCAGCTAGTTTTCAAGGGCCAAAGGCACTCGTAATGAGCAAATTTATGCACAAACACACAGTTTCTACAGTTTCTGATGGAGCTTGGCAAACATTAA
GGGGCCAAATGGCATAAGTAGAAAAACGAATTGAAATAAGAACTAACCACGATGCTCTCGTTTTTCTTATTTCCTATCCCCCCATGTGGGCGGTACATGTTATTGACGCCCA
AAGTGCCGCCCCCATAGACCCATGCCACCGTGCCACGCGCCCACTCTTTTTCTTTTGCTGGCTCCAAAGTTTGAACGGCATCTGCAGAGAAACGGGCTTATTTGACTGAGAA
GGAGGTAGGAGGTCGGCGAAGGACGCATTTATTTGAACTTATGTACACAAGTGCACATATATACATATATATATATACATACATACTGCGCTGTGTATGCCTGCCATATTCA
TGTTCATGTTGTTTTCGCCTTTTAATGAGAGTTAACTGAGCGGGGGAACTGGAAACTCCGCGGCCAACGATGAATATGAATGCGCCACTCAACTTAATCCCACTCAATCATT
ATGTGGCAATCCGCCCAGTTCTTTTTGACCATGGGTCTGCAATTGTCCTCTCGGCTCGTACTCGTACACAAGCCCCAACATTATTTGCACCGAACGTATCTCTGTACCGATC
TGAATTCGTATTGCATGTCGAACTCGACACTTTATGGTGCTCCGAAATTGTTTTGTAAGCGAAATAATGGCAAGTTATGACACCCAGTGTTACACACAATTGACATGGTTAT
TGATTTTCAATTATATAAATATTTGCAACTAAGGAAAAAAGCCTCTGCTGCCACTGCTGCCACTGATGTCCTGTGGAAAAAGCATTTAAATACCGTGTTCATCGGCAATCTG
TGTGCGCTTCTGTGTATTTTCAATTTGCATACACCTGCCCATGAAGAGGCGAAAGTGATGTGTTAAATGAAAATATGCTCCTAATAAACGTTTTTCATCGGTTTCCCGTATA
ATTGTTTTCTGGCTTTCAACACTACACTGCACGAAATTTAAGGCTAACGCGCGAGTGGGGAGCGTGTGGCATGAAATCAATTTCATTCGTTGGTTTTAAAGCCCCGTGTCCC
ATTGATTTCGAGGCAAAAGCTTATATTTGTCAAGATCATAATCAAAGCTGTCGGCCATTAAATAAAATGCTCAACGAGAGAAAAGAAACCCATTAAAAACTGAGCGAATTAA
AGCGATATAAAGCACCATTTTTTGCGTGTGAGTATTATTTGCGTGGCGAAATGAAATATACTCCTGAGCTGAATTAATGAGACCCGGTAGGCGGGTTGAAGTGGGTGGGCTT
TGCGGACAGGCATCTTAAGGGCGCAAGGACATCGGGAACATGGAACACGGAACACGGAACACGGAGCATGGAGCATGGGCACAAAGGGAATACAAATGAAAGGTGATTTTTC
AGCTCCCTTTGGCCTTTTGAGGATTTCCATGGGGAAACCACAAAGTGCGGTTGACAGGGCCTAATGTGAAATTAACATGGCTTATATGCAACGCATACTGCCCTGTTGCATT
CATGTCAGCCACACACACACACACACACATCCATGCACCCAAACACCGACAGCTCAACCGTCCATCTTTGTGTAGCGGGTCAAGTTGCCATCAAGAACACGGTTGTGCCACT
GCCTCTGTTCATCCTGTTTAACAGCCCGGTATTGGATCCCGAATACCAGCCACCACCCTCCGCTCCCTACAACTTGCCGCAAATTTCCTTGGACCCAATTTATTTTGCAATT
TATCGCTGGATAAAAAGTGACTGCGAAAAGCATTTCAAAGCGGAGGCGCAGGACATTGAAAAAACAGTGCCATGACAACAATTACGGCGTGTTTAAGTAGCCCGTCTGGGTG
AAAGTCACGCCGTGAACTGCCACTTTCTCCCCTTCGTTGGATGGCATAAAAGATAGATTGTTTATAGTTGGTATCAAAAGGAAAAAGGCCTGGCCAATGTCGCAGATTGAAT
TGAGTGCTCTGGAGGGAATCGGGGTCACCAAATGCTGGCCTCAGAGTCCAGAACGCCAGCTCTAAGACTTGTATTATTATGGTATTAAATGGTCAGGGACACCAGAGTACAA
CTTTAAAGTATCTTGGTATAAAGTAAAGTATGCAGTTTTAATCCTAAACTTGGCTCAAATGACCTTATATTCCACCATTTTTAAGTAACAGTCCAGTAGATTTCAATGCATC
ACAATTAGAAATGAATTCAGCCCTGAAGGCTGAAAAATAAACCCAATTAATAATTTAAAGCTTCGGCTGAACGATTAGCATGCATTATTAACACTCACAAATATATGCCATA
ATATGTACTTTTATGTAGAGCACTCTCAAATTGTGTCACATTCAGCACATTCCTTTCAGTTCGTTTAATTAAAGTGTCCAATAAATTAACGCTCGACAAGAAGGACAAACTT
AAATGCACTTCAAATGGCGCAGAACCCATCAAGACTTTCCATCAATTTGCAAAAGTGTTGGCCAAAAGTGAGGCTCCCCACCTGGGAACTTTGTTTCCGTGAACTTTGCACC
GGCACGCGGTGAACGCGAAATGCATTATTCACTTTTATGTCTTTTTCCTCAACTACTGTCCGGGACAAAAAAATCAGTTTGCAACTGAAGAATCTTACCTCGACTCGAACTC
CAATTGCGATGTCCTTGTTGCCCAGCTCTCAGCTCCTTGATGATGCAGCGCAGCGGCAGTTGGGCTCTCCGTAAATCGAGCGACATAAAGCGGTCACCCAACATTGGCCATT
CGATGGGAGTGTGACAAAGCTCCCAGCTAAACTTCCCCGGCCATGTCATGCATGCGAAATGGAGTTCTCGAGAAGCAGCTAGTGAATCTTTAATGGATTCATCGATTGCGAA
GCGAAGGAAGCTGCAAAACTTTCGGTTTTTAAGCCACGCTCGGTTTGGTAGTTGGACACTCGATAAATTTACCCAATCACCAGAAAAATATGCAACAGCTTTTGGGCTTCAT
TCTACTCTAGTTTTCCAATTAAAATGCAAAGTTCTAGAGTTTCCAAAACTCAATGCTCCTGGACCCCCATCCTCGCCGGCGTCGTCATCCTTGGGAATGGAGATCTCACTCG
ATTTGGCCATAGAGTGTGTAGGACATTTTAATTTGTTGTGTTTTACAATTTGCCCCCATACAAATTCATGTCGATTGAGGCTTTTTCGCCTGAGTCCGCTCAGTTTATGGCA
CTTTTATAACTCACATGTATATATGTATATATGTATGTACAGACAGGCGAGGACCCCCGGGGCGGTAAAAAGTTGCTCTTTTTTTTCGATGGCCTTAACTTGGCGTATCAAT
ATGTGGGGGGAAAACAATGGAAGCAATTTTTTCAAATAATTTCACAGGCGAAATCAAAAGGAAGCAGCCCGGGAAGAACAAGAACGCGCGGCACATCGACCATCTGATAC
CTCTTGTCTATGCTGGTGCGGTGTCCATTGTGGGCGCCACAAGTGGAGGGTAGTGATCTGTTGATACCATCGACATTCATCCGAGAAAAAAGCGTATCCTTGTGCCTTTTCA
ATCTGGCCATCTTGCCATCCGGCCATCCAAGCATCGGCTGCCATCGAGCGTCGAAAAGGCCAAACTTTTGCACAACCGTTTTTTATTGTTGACTCAAATTTGGAATTATTTT
TATTTGAGCACCGTGGAGGGACACACGCACTGAGTCCATAAAAATCCACTTTCAGTTATGTGGCAGTGGGTGTAAGTGGAAATGCCGAAAGATGGACTTATTGTCATTGAGT
ATCATCGCATCGAATGGGGCGCTAAATCAAAAACAACGACCAGAGGGCAATCCGGAATTGGATTGGTTTCGAAATCTGGAATTCGCGGTGGGGGAAATTGAAAATGGGCCA
TATTTTGGCAATCGATTTGCATGCGCCGCTTAAAGTCATTTCCAATTCGATAATCATAATTTATGAAGCTGTTTTGTATTGCAATTTTCGCATATAAATTACATTAAATATT
TACGTCAGCCGTGAAAAATACATAAATAGTGCGTCTCTGGTGTGTCCGCCGATTCCGCGCGAGTCCTTTGTATCCGCATATCCTGTTCAAACAATACTGCGCTAGAGTCGAA
ATCCAAATATTTGCGCAGCAAATGAAAAGCATTTAAATCCGGTCGCAGAGGTTAACCGTAGTAGTTTTTAAGCCGCCTGGCTGCCGCAACCGGCAAACTTTTCGATCCACTC
TCGCGACTAGCGAGCAATTAGCTGCGGTCGCATTTCATTGTGCTACCCGCTATTTGCACTAAACTTTACACTCATTAAACAGGCGAGAAATTATTGCAGCGAAGTGTCCGGG
ATTAAAGTGTTGCATCCCCATCCCGCTGCCAGTCGGATGATGATGAGGCATGATGGGTTTCTGCCAGGTGGTCGCACTTTTAATTGGTATTCAACTTTCTCCTTTGACCCTC
ACAACTCGAAACATGCTCGTAATTTACCAACCAAAATAAGCGAGTTTCCCAGTTCCGTTCCAACTTCCTTGGTCACCTCAAAAGCTGAAGGCGAGAAGCCTGTTCGCTAAGC
TAAGCCGCACTCAAAGGTTGGCTCTGTGTTTGCCTTACCTTTGCAGTTGAGATGAAGATAAACAGAAATAAATGAAATGTTACATTTGTGTACCTTGAAAATATTCAGAAGC
CATGTTTATAAATGAAATTATACATTTATACTATATTTAATGCAGTGCCTGGGAAAACTCTAAGACATTGGTGTTGCATTAAAAATAAAGTGCTTAAAAGTATGCAACACA
GAATTGATAGTACTGCATTAGCATGTTGATTCATCTGAATATTGAGTTACATACCCCACATGTTGTACTTATCAGATAATTATTTGATTTACATGCTTAAATCATATCAGTG
CTAGCTTACCCAGAAGCCCTAAATTACTTCTATCCCAAGACATTCTCTCAGTGTACATGCGGTGGGTATTCCCAAAAGTTTAGGAAAATTGCCCGTGGCATTTGGCTTGCCA
GAAACGGATGCCTGTCGAAGTGCAGCTAGTTCTTTTGGGCAAAGGCCAGTTATACTTTTGCGGTTGTTAGCTTAATTTAATGTAACTTTCGACATGGTCAGATGGAGTCGAA
AGGTGAGCGGTGTATCTGGAGAAGATCGAAAAGCAAATAACCAGCTGGAAAATGTTTGGCCACAACCCTCTGAGAATCCACTTGGCCTTACGAGTGGATGGATGGTTGATGG
GTGGAAACCGGAGCACACGCACAGCACAATTAGCAGTTATGAAAAGTTGAAAGTTCATTTTGTGCAGCAACAAAACGAAAATGTGATGAAAAGTTCGCACTGACCTTTTGGA
ATTTTTCATTGTGAACGAAGGAGCTGGCCGAAAATGATTTTGCCAAAAAGCCGTGGGCGACAATAAACAGATGAAGTCTGAGGCATGTTTGCGGTCCACTTAATTGCAGGCC
AAGAAGCTCATGAAATACCCTGGTAGCCACAAACTGGTTGCTCCATCTGCAGACCAAACGAAATTATGGGATTCATTCGCAGAGGGAATGTGCAACAAAGCAACTAAAATGT

FIGURE SHEET 30

GAGATTTTCCATTTGAAAGCCCACACTTTCGTTCGATTTGCCACAATTTGCAGCATGTATAAAGTTATATCCGTACTGCCATGATATTTCTGCATGGCAGAACCATTTCAAG
TTTGTCACTCGAACACGCACCCAAGTGCAGAATGCCGGCATGTGAGGATGCGAGGATGCGAAAATGCCAGGACACGCGGAGGTGGCCGGTGGCTTTTCGATAGCGGAGGATG
GTCAAGCCACGAGGGGCACCGAAACGGGAACCCGGCTCATTGTCATTTTCCGACCCGAGTTGGCTTTGAGCACCTGAGAACCAGTCGACCCAGTCGACCCAGTCGACCCAGT
CGACCAGTCTACCTGTCTACCTGTCGACCAAATGTGGACACAGGCTCCATTCGAGATGGCTGCTCGTCAACAGAGCCATCAAATTGTCATTTTCCGAACAAACGCGTTGGTT
TTCACTCGAAATTGGCCGAAAAATAAATGTCCTTGTAAATCACTCGACTTTTATCTTTAAAATATGATTTAAATTCAGCGAATAAAATACCCAATAGCTGATAATTTATGCC
CAATTTATTATGAATTAAAAAAAAAAAAAAAGATATGAAGATTTGATATTAAAATGCCCAAAAATACCATATATATTACCAATTTTATTGATTAACATTTGCTCATTTCAGTC
ATAAACAAATCCTTTGATAAAGCCAATAAAATAATATTTTAAAAAATTTTAAACCAAACCCAAATCATAATATTAATCTTATGCACACCAAACTAAATGTTGAATTGGCAAG
ACGAGAATGCTAATAGTACATAAATCACAACTTATCACTTTGGGAAACAACCATTAAAAATGCTATTTAAAAGCAGACAAACAGCGTAGGATGGCTGGTACATTTTTACTGG
GATTTTGCTCTCCTTTGGCAATCTGTCATGTTTTTGAAGTAGGTCCTCGTCCTCGACAGCATCGTCTGCCAGATAAGAAATGCATATTTAATAATTTTATTCTTTCACTCAG
CAGGGGCCCTAGGAGCTTAGCTTGTTTTATGACTGCGACAAAGTAATCCTGTGCAAAAGGTAGGAAGTGGGAAATCACTGGTTCCAAAGGGGCAGATTATCAGGGATACAAC
AGAGATGTGCTCGACTTAGCTGGCATTATGAGAAGCTCGTGTGGAAAAATTGGCCCCAACCAACACACAACACAAAAAATGGGCGAAGCCAAAGTTCTCGATGTCGATGTTGT
TTCCCTTCAATCAGCGATTCAATTTACGAGTGCCAAAATTCGACACAGTTTTCCCCTTAGCCATTTTGCCACTTGCCACTTGCCACTTTTCACTGCCGGCTGACACGCCCCA
TCCGGAATGCCGCGAAAATGGGCGTGCGAATATTTTGCCTTTCGCTGAATTCATTGTTGGCCCATGGCCCACGTTGCGGTCCGGTTTCTATTTTCAGTATATCCATAAAATT
AACATTTTTGTTGCGGACTCTCAAAATCACTGAGCTATGGTCCTCGTGGTCCTCCAGTTTGCATTTCCCAGCGAAAAGGGAAAGGGGCAAGAGGTAAAGTGGCAACGGGGTA
GAGGGTGCGTATATGGCGGCGCCACGCGACGGAGCTAAATTTAATTTTGTTAATGAGCTCGGCCCCTTTCCGTTTCATCATCCCGCCAATGGGGGCAGCCAGCCAGCCAGCC
AGCATGGCGCAGCCACACTCTGCCGTCCTATTATTTTGCATTCGCAATTCTAAAAGCCTCAACCCGTTTGTGAGGTGGTCCGTGGTTCCAGAATCCACTTGCCAAGTGGAGC
GAGTGTGTGTGCGTGTGTTTTGGCCTTTATTGCGCTCTTGGTGCCCGCCAGGATTAGACTGTGGATAGGCATCGGGATTGGGATTGGGTTGGGGATGGGGAATGGGGAATGG
GAATGGGAATAGGGGATTGCCATTGGGATGGCGATTAGAACGACAATGGACCATTGCCGGAAACACTTGAAAACTGTTTCGTGTTGGCTGGGGACTTTGATTAAAAGCGGCA
CCATTGCTATCATAGCGGCAAGCCGAAAAGCAGAAAACAAAAACAAAACAAAAAAAATGAGGGAAAATAAACAAGTTGGATCTGACAAAGACATTAACATTTTAATTGTTCC
GGCAACGCAATGGAGTCGTGGACCAAGCGTAACAATTGCCAATGAGTGAGCTCGAATAAAAGTTTATGGGGTGGAGGCGAAGGGCAGTCGGGAAAATGGCCATTAATATTGG
AGCCATAAAAGTCAAAACTTTTCGACGGGTGGCAAATGGGAGGCGTGGATTAGCAATATATCCCATGCCGCCAGGAAATTGACCCTAGTCATTATTTGCCACAGAGATGATT
ATGTTAGGCAATAAACAAACTTAAGTTTTGAATCTGCTCGATTTGTGAATGAAAACGCCAAAAATATACATATATAATCTAATGTAAATATTTTAGTTATGGATTTTGAATG
GCTTGGATTTTTGTATTGGTTGGTAACTTCTCTGTTTTTCCAATTCATTGCAGCTGCCTCGGTGGCCGTTTCCTCCTGGACACTGGTGGCCATCTCTTGCGAGCGCTACTAC
GCCATTTGCCACCCGCTGAGGTCGCGCACCTGGCAGACGATCAACCACGCCAACAAGATCATCGCCATCATCTGGCTGGGCAGTCTGGTGTGCATGACGCCCATCGCCGCCT
TTAGCCAGCTGATGCCGACCAGCCGACCAGGTGAGTGTGCGCCAGGATAACACTAATTCAGGAACTAACCCCCTGTCTGGTCAGGACTTCGCAAGTGCCGCGAGCAGTGGCC
GGCGGATAGCCTCAACTACGAGCGGGCATACAACCTGTTCCTGGACCTGGCCCTGCTGGTCCTTCCTCTGCTGGCTCTGAGCTTCACCTATCTTTTTATCACCCGCACCCTG
TACGTGAGCATGCGCAACGAGCGAGCCATGAACTTTGGCAGCAGTGGGCCGGAGGTCACCACCTCCTCCTCTGCCGCTGTCGCAGAAGCCGGTAGCCAGCGACGCGCCAACG
GAAGCCATTGTCAGTCCCTCGACACGATTGTGCCACACCAGCACAATCCTCACCAGCAGCACCACCACCATTCCCAGTACTACTATGGTAAGGGGTTTATGCTTATTTTTG
AGATATTCGTAGATTTTTTTGTCAGCCAGCTGGGTCAAAGGGAATTCTGGGTTCAGAGGGTTCAGTGGCGTAGCATTAATAGTTGCGACTGTTGCCGGGGGTTGGCAGGGG
GCGTCATTCGGGATTAGCTCAACTACTGAGCTTATCCGCGGACCCAAAGCCCCCGAACTGCGTCACACACAATTTGCACGGCAAAGTTTCCCACGCACGAAAAGATTCGTGT
GTCTACGACTCCTGCCTAAATCGCGGAAGTGAATTGAATGGCTGGCTGGGGGATGGGGAAAGGGGGCTAAAGCCTGGAGGAGCACTCTGTCCAAGTTTTTCAAGCGTGTGTA
CGAAGAGACAATAGCAAGCGAAAGCTCTCCCGGCGTTTTTACTTTCATTTGCATTCGCATCGCCGGCTTTAAGAGCACCACCCATGGCTAGATGGCCCCTTCTCCCATCTAC
GGCGTCCTCCATCCTCTCCAGCTGCGGATGCTGTGAATAATCCCCAGATTATGGGCACAAAAATTCTAGGCAGCCCAAGGCAAAGCCGGCGCAAATCAATTCTGCTAATTCG
CCTTTAATGCCGCCTTAAAGCGTGCTCCCAATATATACATATCTCTATGTAAATTTAGAATATATTGTAGTATATGCTCGTCTGCAGCAGACATGCTTGACTGTAAATATAG
AACAGTCCCAGTGGAGAGGCCTGCTGAAGAGGAACCTCCTCGGTGCTCCACCATGGATGCGAGTTGATAACGTGACGTGTATGATAAATGCACACCCACAAAGCGAAGTCAA
CCGTAGAGTCCCCGTGGAAATCACAGTGAGCCTACCGAGCAGCGAACTTACTTTAATCTTCGAGAATATTACTCTGCCTCTGGTCCAATTAATGTTGAGAAGAGGAGCTCCT
GCCACGGAAAGATGGAAGCACTGAGGCAATGGAGCAGTGGAGCAGTGGAGCAGTGAGTCCCTGGGCGACTCATTTGCATGAAAGGGTTGGCGCCTCCAGGTGGTAAGTCCTT
TATGCGAAATGGAAAGTGTGCATTAGCACCTTCACTTGCGCCGCACTGAATTATCGCAGCACTCGAGAATTCATCGAGCGATTCATTTCCTAGGCAACTGACCCCGCTTGAC
CCATGGACCAACACGGACATGCCAGTCCACCCGCTGCTGAAAAAGGTCAATTTAAATGCCGTAAATTATGCACGAGACGAGCATTTAAGCGCTCGTTAAAGGCCGTTATCCC
ATTGCCAGGATTCTCCACTGGAGCCCGTCGGACGCGTGGCACTCATTTTCACACATACTCGCACTCCGGTCGCCTTTCTGTGTGACAATAAATTTGATTAAATTTTTAACTT
TTTCGTTCCGTTTAAGGCCATCTCGTTTGACATAACCCCGTCATAAATTTCCAAGCCAAAGAGCATTTATGTGTGTTGGCTGGATCGCAAATGTCCCATCGCCGAACAAGCT
GAACATTTTCATGGAAATTGCCATATTATTAGCAGCTTCACATGGCCCCCACTATTCGTGCCAAGAATCCGCATGCTATCCATCCCCAAACTGCGCGCTTGATGGCAATTTA
AATTATGCTCAGCTGAACTACTCGCCCAACCGCAAACGGCAATAAAATTGCGAGGTATATGGTCTCATTCCGCACTTAAGTATCCATTTCCCCCTTACATGTATCCCAAATG
ATTCCCATCACATGCTGGTAGCTACAAATATAGTATCCTTTCCATGTAACATCTACAATGTAGCACAAATTTTGTGATATATGGATATAATCCCTTCATTGCATCATTAATA
TCTACGTGTCCTTCTCTTGAAACTCAGCCAAATGCCATTCCTGCTGGCGGTGAGGAAGGTGAGCGGAGACACCTCCACCCTGCACTGATCCTCATCCGCCAGCTGGACGCAG
TGGTCCTGCAGTAGTGTGGTCAGTGCAGCCTTCGTCTGGAGCAGGCCAAGTCGCATTCCCGGACACATTCGTGGTCCAGCTCCGAAGCCCAGAAACCGACAGCCCATTGGTG
GCTGATTTAGAAAGCGTTCTGGACGAAACTGGTTGGGTGCCGGATACAATGCGGGATCTCTGAAATCGAAAAAGACTGTAGTTCGTAGACAATCTGAAGTCTCACTGAAGTG
TCACTCACAAATGAATGGCCTGAACCGGCAACACCAGCACGGTGCCCAAATGCACCTTAAGTTCGCTGCTCGCTCCTGATTTCTGATCAGGAAGGGTGAATGTCTTGGTGCA
GCGCTTCTGCAGAGCCTGCATGGCCGGATGAAGACGCAGTGCCTCCAAGAGCGCAGCTTCGCTATAGCGAAGTTCACCCAGAGCCACTGGATCAATCAGATTGCCAGCGTGG
CGCTGGGCCACTTCATCCAATTCGATGTGTAACCGTCGCTGTGCATCCTCATTGAGGGCCAATTCGTAGAGGGCGAAGGCGAGTAGCATCGCCGAGGTTTCGTAGCCCTCCA
GTAGGAGAGTGGTGGCATGTCCAGCTAGCTCCTCCTTTCCCAAGCCTCTTTTGCTTTCCGCCAGCCATTGCAAGAGGTTGTCGCCACCACTTCGGGCTTCCACCAATTCCCT
GAACCAGTGCTGCAGGGGCAGGGGTACGTATCTGTATAAGTGGATAGTGGATCAAAGGTTACACATTGGCGAGTTCCACATTACCAACCGGTGGCCAATGAGTCTACCCAAT
CGAGGCGTGTGCAACAGTGACATGGTCTCCAGCAAACTCCAAACGCTTGGCTGAAAGAGCGGAGCGAGCCATTCCAGCCAGCGACTTGGTTCATGAGCCACCCTCATGTGAA
TCCCCAGGCAATGTGCATCCAGACCGAAGATCGCGGAGGCAACCACTTGTAGCGTATAACGCGTGGCCAAGTCCTTGGCCTCGAAGCGTCCAAGGGGCACCTGATCCCGGAG
CAACTGACATGCTCTGGCCACATGCGGCAAGGTCTGACGCACTCGGCTGGGTGTGAAGAGCGGCACCAGGTCGGCCCGCAAGACTCGCCAGCGGTCTCCGTTGGCGATGAAG
GGATTGTGGCTAGCCAGGACATCGCGCTGGTGGTCCACATAAACGGCGTTGTCCGCACAGTCGGCAAAGTTACGGCCCACCAGGATCTGGCGCAGCAACTCCTGGTCACGCA
CCAGAATGGCGGGTTCGTTAAACAGGCGATAGAAACCCACATATTTCAGGCCCGGATTCCGACTGTCAAGAAAAGTTCATTGCACGTGGCCTTGAAGGACTCTTAAGATTAG
TGTATCTTACGTATAGATTTCCTGGTAAATCTCTCCATATGAGCGGCGGCCCAGGGCAAAGTCCAACATATTCCCAACCAATGGCCAGCCAAAAGGTGCCTCCAGTCCCAGG
CGACGCCAGTGATTGCCCTGCCATTTCCAAAACAGGAGCGTGGCGGCCAATAGGATAAATAGCACCAGTGGCAGCATTATCGATGAGCATCTGAAAAGCTTGAAAGGCACAA
TCCCACTCAACTTGCATTCCCAGTTGAGTTGAATATCTCAAAGCAGCTGAGAGTGAACCACGTTGCGTGATTCCATGATTGCGCCTCATTAGCCACGTAATTGCTGTATATC
AATCAACTGACGGGTATGTTTCAATCACGCAAGATTCGAGACAGCTAGCATTGGGCTTTGAATTAATTTCTATAACCACTATGACTCAATATGTTTTCGGATAATAATCTTA
GTATTTCAACTTATATGCATTGTACACATCCATTCAGTTTTAATTTATTTAGTGCATTTCTTCTGTTCAAAATTTTGTTTTTTTTTTCTGTGCTTAACATTCAATTTGACAA
TCGTATCCATTTGCGAGTCAAATCAGATGTGTTTTGTGGATACCCTGTAACTTTCCATCTTGCAGATTACGGCCACTGTGGCAGCAAGCGGCGATTGATCAGCGGAGGAGGA
CCCTGCGAAGGAAGGAGGCATCTATACTGCATGCGGAGCGCATCCGTGAAGTCCCTGCGCCATCAGCAGATCAACGGAGGAGGTGGAACCCTAAGTGGAACTGGAGCTGGAA
ATGGCGAGTGCTGCAGCCGGGTGCACAGGATGCGCCAACAGATGCAGCTGCAGCAGCAAGGCTACGTGAGCGACAATGAATCCCGGCGCAAGTCATTGTCGCAGCCCAGTCT
GCGTATCACGGAAGCGGGACTGCGGAGGTGAGTGCTTCAGGGTATTCTCATCCACATATGACCGACCGATTCGTGTGTCGATTTTCGCTTTCACGGTACCTTGCAGATCCAA
CGAAACCAAGAGTCTGGAGAGCAAGAAGCGTGTGGTCAAGATGCTGTTCGTCCTAGTGCTGGAGTTCTTCATCTGCTGGACGCCGCTGTATGTGATCAACACGATGACCATG
CTCCTTGGGCCGACGGTGTACGAGTACGTGGCTATACCTCCATCAGTTTCCTCCAGCTGCTGGCCTACTCATCCAGCTGCTGCAATCCGATCACCTACTGCTTCATGAACG
CCAGCTTCCGGCGCGCCTTCGTGGACACCTTTAAGGGGATGCGGGTGTGCGAGCGCCTGTGCGCTCCCTGCTGCTTCTGGCGGCGACGCTCCAAGAACGAGACGAATCTTTC
GGTGGCCGGTAACTCCATTGCGCTGGCCAACTCGGTCATGTCCAGCCACACGATCCTCGAGAGTCCGCGACTCTGAGCCAGCCGCCGCCGACGGCAGATGGCCGAGTCGGAG
TCACTTTAGGTTAGCAATGGGTACGGATCGGGAGCGCTAATATGTGATGATGCAATGTAACTGCAATAGAAACTGTGGTTGTCAACGACATTCCGCCGCGTTCTGAATGTGT
GTGGAAGTGGCCAGATGGATAGGTGGGTGGATGGAGGAGGTGCCATGGCAATTGCCCCTCCAAATTGCTCTTCATCTTACCGATAAGTATTTTAATATATGCGTATGTGTGT
ATGCTTAGAGTTGCGCTAAAGCGCTTTATTCCGTGTGTGTTATTCGACTTTGACCTCGCGTTCCATGTGTTTACACCTTAAGGTGGGTGCAACCAGTCATGCCAATCTAAAT
TGAGATCGCCACACGCTATGGGGAATGGAAAGTGCTCAGCTGGATTAAGATTTTTAAACAAACAATAGGCTAAGAAATCAGCACAATGGTTTCACGACATGGTCACAGTATG
TCGTGAGCCAACAAAATGACTTCCAAATGGGGGGCGCTCCATAAAATTGTTTTAATTTATGGAAAATGGTATATGTGTTGGAAGGATTTGTTGGTTCAAATCTGGAGTTGGG
TCAAAAGTTGGTAAACAACTCATTTGTCACACATGTCGTATGAGTGATATTTACGCAAGCTATCCGACCATTGTGCTGCACTTTCTAATTTTATTTCCACACTTTACGTTCC

AAAGATCAATCCCAAAAGTAAAAGGTATTCTGATCACGGTGCAGTGAAAGGTTTCAATGAATCAAGAAGCATTTGACAGCAAGCTGATTTTACAGGGTACACAGGGAAGTCC
CACTCTCAAAAGGATTCCCATCATTCAAAAATGGTTGTCAGACGGAGCTTATGGAGTGGCAGACACTCCGAGGTCCAAAACTAAATTAAGAAACCCAATCGGGTGCCAAAAT
CCAATTTTAAGTAGACGAAAGTAGCTGGAAATGCCAATCGGATATATTCGGGCGGGATTAGCATTGTG
(SEQ ID NO: 61)

Exon: 89..105
Exon: 176..924
Exon: 1025..1252
Exon: 12710..12910
Exon: 12965..13303
Exon: 17202..17499
Exon: 17579..17996
Start ATG: 89

Transcript No. : CT21155
ATGTTGCCGCGCCTGTGCGCCGACGCTTGCCGTCAGTGCTTCGCGAAGATTGCGAGGCGCGACACACATCGCGGAACCCGAACTCCATACGGCTGTGCCGACACTCAATCCC
GACCTAAGCCCAATTTTTGTTGCGTGAAGTTGACGAAGTGTGTTGTACTGCTGCAAGTGCAAGTCCCCGATTGCTAGTGCTATTCCGGGATCACAAACGAGCATCCTTCTT
CGGCCTCACGATTGACGCCTTCTACCACTACCTGCGCCAGGCGCTGCCGCTGGCCAAGGAGGCTGCGATCCACCTGAATGCCAGCAACGAGATCTCCGCTGTCGGTGACGGG
GTCACCATCACAGGGACACCCGGTGATCTTCTCAACTACAGCGGGCTGGAATTGGATCTCGGACTGGATTTGGATCTTAATCTGGACATGGACCTAGCCACTACTCCCAGCA
GCAGCACCTTGGCGCCAGCTGTAACCGTCAGGACACCGGGGAACAGGAGCGTGGTGAGGGTCAGTGCCGATGTGCCCATCTGGGTGGTGCCGTGCTATTCGGCCATCCTTCT
GTGCGCCGTGGTCGGGAACCTGCTGGTGGTGCTAACGCTCGTCCAGAATCGTCGCATGCGCACCATCACGAACGTTTTTCTGCTCAACCTGGCCATCTCGGACATCCTGCTC
GGCGTATTCTGCATGCCGGTGACCCTGGTCGGCACTCTGCTAAGGCATTTCATATTCGGCGAGCTGCTCTGCAAGCTCATCCAGTTTGCTCAGGTCCCTATCCCCGTGACCG
AGCCCCATAATCGCTTGTCAAATGGCCACGAACCAGCACGACCTCGAGCGGATTCGCCAGCAGATCGTGTTGGCCAACATTCAGGAGCTGATAAAGAAGATGACACGTCGCT
GCTTCGACGTATGTATCGCTATGCCGGAAATGGAGTTGCGCTCCACGGAGCGCGACTGCCTGGCCAACTGTATGGATCGATTCATGGACTCGGTTCAGCTGCCTCGGTGGCC
GTTTCCTCCTGGACACTGGTGGCCATCTCTTGCGAGCGCTACTACGCCATTTGCCACCCGCTGAGGTCGCGCACCTGGCAGACGATCAACCACGCCAACAAGATCATCGCCA
TCATCTGGCTGGGCAGTCTGGTGTGCATGACGCCCATCGCCGCCTTTAGCCAGCTGATGCCGACCAGCCGACCAGGACTTCGCAAGTGCCGCGAGCAGTGGCCGGCGGATAG
CCTCAACTACGAGCGGGCATACAACCTGTTCCTGGACCTGGCCCTGCTGGTCCTTCCTCTGCTGGCTCTGAGCTTCACCTATCTTTTTATCACCCGCACCCTGTACGTGAGC
ATGCGCAACGAGCGAGCCATGAACTTTGGCAGCAGTGGGCCGGAGGTCACCACCTCCTCCTCTGCCGCTGTCGCAGAAGCCGGTAGCCAGCGACGCGCCAACGGAAGCCATT
GTCAGTCCCTCGACACGATTGTGCCACACCAGCACAATCCTCACCAGCAGCACCACCACCATTCCCAGTACTACTATGATTACGGCCACTGTGGCAGCAAGCGGCGATTGAT
CAGCGGAGGAGGACCCTGCGAAGGAAGGAGGCATCTATACTGCATGCGGAGCGCATCCGTGAAGTCCCTGCGCCATCAGCAGATCAACGGAGGAGGTGGAACCCTAAGTGGA
ACTGGAGCTGGAAATGGCGAGTGCTGCAGCCGGGTGCACAGGATGCGCCAACAGATGCAGCTGCAGCAGCAAGGCTACGTGAGCGACAATGAATCCCGGCGCAAGTCATTGT
CGCAGCCCAGTCTGCGTATCACGGAAGCGGGACTGCGGAGATCCAACGAAACCAAGAGTCTGGAGAGCAAGAAGCGTGTGGTCAAGATGCTGTTCGTCCTAGTGCTGGAGTT
CTTCATCTGCTGGACGCCGCTGTATGTGATCAACACGATGACCATGCTCCTTGGGCCGACGGTGTACGAGTACGTGGGCTATACCTCCATCAGTTTCCTCCAGCTGCTGGCC
TACTCATCCAGCTGCTGCAATCCGATCACCTACTGCTTCATGAACGCCAGCTTCCGGCGCGCCTTCGTGGACACCTTTAAGGGGATGCGGGTGTGCGAGCGCCTGTGCGCTC
CCTGCTGCTTCTGGCGGCGACGCTCCAAGAACGAGACGAATCTTTCGGTGGCCGGTAACTCCATTGCGCTGGCCAACTCGGTCATGTCCAGCCACACGATCCTCGAGAGTCC
GCGACTCTGA
(SEQ ID NO: 62)

Start ATG: 1

MLPRLCADACRQCFAKIARRDTHRGTRTPYGCADTQSRPKPNFLLREVDEVCCTAASASPRLLVLFRDHKRASFFGLTIDAFYHYLRQALPLAKEAAIHLNASNEISAVGDG
VTITGTPGDLLNYSGLELDLGLDLDLNLDMDLATTPSSSTLAPAVTVRTPGNRSVVRVSADVPIWVVPCYSAILLCAVVGNLLVVLTLVQNRRMRTITNVFLLNLAISDILL
GVFCMPVTLVGTLLRHFIFGELLCKLIQFAQVPIPVTEPHNRLSNGHEPARPRADSPADRVGQHSGADKEDDTSLLRRMYRYAGNGVALHGARLPGQLYGSIHGLGSAASVA
VSSWTLVAISCERYYAICHPLRSRTWQTINHANKIIAIIWLGSLVCMTPIAAFSQLMPTSRPGLRKCREQWPADSLNYERAYNLFLDLALLVLPLLALSFTYLFITRTLYVS
MRNERAMNFGSSGPEVTTSSSAAVAEAGSQRRANGSHCQSLDTIVPHQHNPHQQHHHHSQYYYDYGHCGSKRRLISGGGPCEGRRHLYCMRSASVKSLRHQQINGGGGTLSG
TGAGNGECCSRVHRMRQQMQLQQQGYVSDNESRRKSLSQPSLRITEAGLRRSNETKSLESKKRVVKMLFVLVLEFFICWTPLYVINTMTMLLGPTVYEYVGYTSISFLQLLA
YSSSCCNPITYCFMNASFRRAFVDTFKGMRVCERLCAPCCFWRRRSKNETNLSVAGNSIALANSVMSSHTILESPRL*
(SEQ ID NO: 63)

Name: GASTRIN/CHOLECYSTOKININ RECEPTOR like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384801
AGCTGTAAAGTACTACTACTACATATGTGTACATTCAAGTTTCGCTTACGAATGCAACGTTCCATTAAAGAAGCCACCAACTTTAAGATACACGCATATAACAACTCCTGAG
AATTGTGAAAACTGATCTACTTTGATGATCAACTTTCATTGCTGCTGATTGAATAATTGAGTAATGAGTAATTGGTATGCTGATAAACAGGAAAACCAAAATGTCAACAGTT
GATAAGGAACATTAAAACACAAATTAAGAGGACGGTAAGTCCTTCACTCGGACCTTTTCAGCTGAGCACCGTAAAGTGCTGAAGGGATTGCCCCCACACCCAGACCGCGCCC
ACTTCCACTTCCACTTTAAAATAATCATCATGGCCCTCTTTTTTGGCCAGGCTGCGGTGGCTTTGGGCCTTAATTGTCGACTCAGTTGCTGTGCTGTGCTGTGTACTCGACA
AGGCGACGAACCGAGAGCCTTGCAGCTCCATAATGACAGGCGGATGGACTCGAACACTCGGACCCGGACCCGGACCCGAACCCGGACCCAGTCCACAGCCAGGACATAAGCG
GCTTGTGCTCCACTCAACTGTCGCCTTTCGGAGCTGCTCCTAGCTTGCAATTAAAATACCCACTTGGGCCCTATAAATTGTGAACGTAGGAAGGCAGGCGGAAAAAAAAAAA
AATATTCCAGCACTTTGGTGTATGGAGTGGGGTCCACTTAATTCCCAGGTACTCACCCAAAAGGATTCCAACTATAAAGACATACAGCGAGAGTGCTTAAGTTTTACTCTGT
TTATTTGCAGCAACCTTAGGAAGTTCAATTGAAACACCAATAGACTAGCGTCTGAGTTATTTTACGAGTACGATCATAATCACAAGTAGGTAAAAAAAATCCATATTCGCCT
TACTTAAAACAAAAAAAAAAACCACACGCAGATCGCTTATCATACATATTTACACCGCATCTCGTTGATCCTGAATGGGTCCTTGGAGGCGAATCCTCTACAGCTTAGAGCT
GAGGACTGTTGACGGCGGCGGCGGCACTATTGGTGGTCACCACCGCATTGAGAAACTCGTGGCGTGAGCGACTCCGCCAGGTGCCCATACCCATGGCTAGGCCGGGATTGAG
ACTAATGTTGGTGTTGGCACTCGCATAGGCGCCCGGGCCTGCGCCCGCCTGGCTGGCGGAGAGTCCTCCACCAGCAGCACCACCGACGCCTCCGCTGGCACCTGCTCCACGA
CGCCAGGGCAGACCCTTGAAGGTGTCGACAAAGGCGCGCCGGAAGCTGGCGTTCATGAAGCAGTAGGTGATCGGATTGCAGCAGCTGGATGAGTAGGCCAGCAGCTGGAGGA
AACTGATGGCCGTGTAGTCGACATACTCGTACACCACCGGTCCGATCAGCATGACCATCGTGTTGATCACGTACAGCGGAGTCCAGCAGATGAAAAACTCCAGCACCAGGAC
GAACAGCATCTTGACCACACGCTTCTTGCTCTCCAGGGTCTTGGCCTCGTTGGACCTGGATAGTGGATTGTGGATTGTGGAGAAATGGGAATAGGGAATAGGATTGGCAGAC
CGGTTATCGGACACGAATCCTACCTGCGAAGTGCCGCATCGTGGACGCGAATGCTGGGCGAGGAGGTCTTGGCCAGCGTCACCGTGGTCGTCGTGGTGGTGGTGATCACAGT
TGGAGCCGTTGGTCTCGTTGTCAAGGTGGTCGTTGCCATATTGGTTGATCCTCCGCCTGCGGATCCCTCTGAATTTCCATTATTATTGTTACTATTTTCTATGTGAAAAAAA
AAAAACATGTGAACACAAGCCAAATATAATCAAACAAATTAACGTTATTAGCAGCAACAAATAAAACAGTTGCCATAACTAATCTAAACTAATACATAGACGTAAAATAATT
AAATTTTCAATTTGGTTTTAAGAGTCTCAAAAATAATATAGCATATTATGTTAAGTGGATTTGTACCTTCTTATCGAGCATTTCACCAACAATCTATATCTTTAAAGTATTA
TGCCATATTACATGTCATGTAACATTAAATAACTTATGAGTTATGATATTTAGATGCAAAATAATACTTGAATGCATTCATAAATGCTCTTAAAGTTGTTTGCATATTCATG
GCTTTTAAGGATTGAAAAAGGATTAGTTTCGTGCAGTTACCAACCATTATAGACTGCGGTGGCGGTCAGGACCAGGATGCAGTTACTACTGCTGCTGGTGCCCGGATTCGGT
GCGCTTCCGCCGGCCGTTGTAGCGGAAACAGGCAGCGATTGCTGCAGGATGCGTCCGCTGTCCTTGGCCATGCCTACGTACAGGGTACGCGTGATGAGGATGTAGGCCACGC
AGAGGACGAGAAGCGGCAGGACGAGCAGCAGGAAGTCCAGCAGGATGTTGTAGAAGAGCTCGTATCCCTGGTCGGGCCAAAACTCACGGCACTTGCAGTAGCCTGTAACGTA

AGTTGGTATAACAACCAGACGGTGTCTGCTGGCCAGGTTGCTATGCACTCACCCGGTCGACTGGTGGGTATCAATTGACTAAAGACCGCTATGGGCGTCATGCAGAGGATGC
CGCCCAGCCAGATGAAGCCGATGATCTTGTAGGCGTGACTGATTGTCTGCCAGGATCGCGAGCGCAGTGGATGGCATATCGCGTAGTAGCGCTCACAGGATATGGCCACCAA
GGTCCAGGACGAAACGGCCACGGAGGCGGCTGAAATAAATAAATGAAAGCGGATTAACTTACAACAAACCGAACACAATGGACACACCCAGCAGCTCCTGTCGAGCGCCGTG
TCAAATCGCTTTCATTAGAAGGTCCTTCGCCCGCGTTCGTTGCATTTTAATTACACAATTTGCATATTGGGACCAGACGAACCGTACCAACCAGACGGACGGAGACATCGCT
CTTCGCACTTCTGCCTGTTCTTGTAGTATGACAAACACGATGCACGATGCAATTTGCAGAAAGTGCAATTAAGCAATTCCCCAGCCACTCCTTTTCTTGTTGCCCGGTCTCC
TTTTCATGTAAGCCTGTTAGATGATTTCCCATTGTAAAGCAATTTGTACCCACACCCTGCCCATCCCTCATATTCTTTTGGGCTACCACTATCGATGGGGCTACACAAGTGC
AAGTCAATAGGGTGGACTGCTGTAGGTGAAAAACTCGACACCGAAAGAAAAAGACCCGATTCGAAAAATTGGTCAATTTTAGCGCAATTAAATGAGAAATAATTTCTAGTTT
TTCAATCGATTGTTTTATAGCACCAAATCATATATCTTAACCGTAAAGGCATCAAATAAATATGATAAACTGATCAAATGAAAGCATAATTTGCTTGTATGTATGCGTGTTT
TTGGAATTTATAGCTTACCAAAACACTTGTGAAATTTGCAAATATACATATTTGTTCATACACAATTTAGACTTCAATGCAATAAGCTGTAAGGTCTATGGTTTCAATTATT
AATACTCGGAAATTGTTGTGCTAGCAGATACATTTGTGGTTTGTTAGATATACATATTTGCTTATAAAGATATATATGTATGTACGTATTGCATTCAAAAGCAATAAGCAGT
AAGGTAAATTGGATCAAGGGACGAGTATTTAAGACAGCCTGTATATGTTCCCTTTACTTTAATGCGACACCTCATTCAGGCAATTAAGGTCCACCCTAATGTGCTTATCCCT
TCAGCCATCTGCAGCTGCATTGTATTATATTGGAATGGTATTTTCTCGCAAAATGATATGCAAGTTTTAGATTGATACGTACAATTGCAGTGATTACTACTTTAGTTACGAA
TGATGATAGGTGATGTAGTAGATATATATTTGCTACCAAAGATAGTTACACCTTAGAATCCTAAAAGAAATAAGTGTTTTGTAGTCGCAGTTCGTTTCTCTAGTTCAACACA
TCCAGCCAACCAAGGTCCACCCTAATGTGACTGCACTTCAGCTGCACTGTATTATATTACATTTTAGCTGCACAATTGTTTACTTTCTATTTTAATTATGCATGCAAACTAA
TATGAGCAGTAAAGATGGATTACCTCCACCTGTTCCGTTGGAGTGCATCCCGTCCAGGCAACCAAGGTCCACCCTAATGGGGTAGCCCTTGCCCTTCCAACTGCATTGTCAC
ATTGGTCTCGCCCTCTGTCGCGTTTGTTGAGTTAAACCCCTGAAATTGAAGTTTACTTTTGATTGACGGCGGTGTGTGATGCTCTAAGCCAGAAATGCTAGCATGTGTGCGC
TTAAACTTAAAACAAACAGCAGCATGGCATTTC
(SEQ ID NO: 64)

Exon: 4080..4056
Exon: 2717..2517
Exon: 2454..2169
Exon: 1736..1001
Start ATG: 4080 (Reverse strand: CAT)

Transcript No. : CT21314
ATGCACTCCAACGGAACAGGTGGAGCCGCCTCCGTGGCCGTTTCGTCCTGGACCTTGGTGGCCATATCCTGTGAGCGCTACTACGCGATATGCCATCCACTGCGCTCGCGAT
CCTGGCAGACAATCAGTCACGCCTACAAGATCATCGGCTTCATCTGGCTGGGCGGCATCCTCTGCATGACGCCCATAGCGGTCTTTAGTCAATTGATACCCACCAGTCGACC
GGGCTACTGCAAGTGCCGTGAGTTTTGGCCCGACCAGGGATACGAGCTCTTCTACAACATCCTGCTGGACTTCCTGCTGCTCGTCCTGCCGCTTCTCGTCCTCTGCGTGGCC
TACATCCTCATCACGCGTACCCTGTACGTAGGCATGGCCAAGGACAGCGGACGCATCCTGCAGCAATCGCTGCCTGTTTCCGCTACAACGGCCGGCGGAAGCGCACCGAATC
CGGGCACCAGCAGCAGTAGTAACTGCATCCTGGTCCTGACCGCCACCGCAGTCTATAATGGTTGGCGGAGGATCAACCAATATGGCAACGACCACCTTGACAACGAGACCAA
CGGCTCCAACTGTGATCACCACCACCACGACGACCACGGTGACGCTGGCCAAGACCTCCTCGCCCAGCATTCGCGTCCACGATGCGGCACTTCGCAGGTAGGATTCGTGTCC
GATAACCGGTCTGCCAATCCTATTCCCTATTCCCATTTCTCCACAATCCACAATCCACTATCCAGGTCCAACGAGGCCAAGACCCTGGAGAGCAAGAAGCGTGTGGTCAAGA
TGCTGTTCGTCCTGGTGCTGGAGTTTTTCATCTGCTGGACTCCGCTGTACGTGATCAACACGATGGTCATGCTGATCGGACCGGTGGTGTACGAGTATGTCGACTACACGGC
CATCAGTTTCCTCCAGCTGCTGGCCTACTCATCCAGCTGCTGCAATCCGATCACCTACTGCTTCATGAACGCCAGCTTCCGGCGCGCCTTTGTCGACACCTTCAAGGGTCTG
CCCTGGCGTCGTGGAGCAGGTGCCAGCGGAGGCGTCGGTGGTGCTGCTGGTGGAGGACTCTCCGCCAGCCAGGCGGGCGCAGGCCCGGGCGCCTATGCGAGTGCCAACACCA
ACATTAGTCTCAATCCCGGCCTAGCCATGGGTATGGGCACCTGGCGGAGTCGCTCACGCCACGAGTTTCTCAATGCGGTGGTGACCACCAATAGTGCCGCCGCCGCCGTCAA
CAGTCCTCAGCTCTAA
(SEQ ID NO: 65)

Start ATG: 1 (Reverse strand: CAT)

MHSNGTGGAASVAVSSWTLVAISCERYYAICHPLRSRSWQTISHAYKIIGFIWLGGILCMTPIAVFSQLIPTSRPGYCKCREFWPDQGYELFYNILLDFLLLVLPLLVLCVA
YILITRTLYVGMAKDSGRILQQSLPVSATTAGGSAPNPGTSSSSNCILVLTATAVYNGWRRINQYGNDHLDNETNGSNCDHHHHDDHGDAGQDLLAQHSRPRCGTSQVGFVS
DNRSANPIPYSHFSTIHNPLSRSNEAKTLESKKRVVKMLFVLVLEFFICWTPLYVINTMVMLIGPVVYEYVDYTAISFLQLLAYSSSCCNPITYCFMNASFRRAFVDTFKGL
PWRRGAGASGGVGGAAGGGLSASQAGAGPGAYASANTNISLNPGLAMGMGTWRSRSRHEFLNAVVTTNSAAAAVNSPQL*
(SEQ ID NO: 66)

Name: GASTRIN/CHOLECYSTOKININ TYPE B Receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384546
CAAAATAAATGTATATAATTTCAAAATTGACATTTGCTAATGATAAGCCTCTGCCTCATGCCCATGCTTATGCTGAAATATGGACAAAATTCGGGCACAGAGCAAGAGATGC
GCAGAGCAGCTGCCATCCGGAATGGAATGAGTTGGAATTTGGGCAGCAGATGGCGCTGTCGCAAAAGGAAGGAAGCTAAATTTCACTGGGCCCCAACCAAACGGAAATTTCC
ACCCACGGCTCAGAAATTGAAGCTGTTTATGGTAAATTAAACCAATTGAGTGGGATTACGTTTTGCGCACTACAACAACGCCACCAACCTCCGCATCAACTCCGCAATCTCC
GACACCCCGAAGTGTTTCCAACTGGGTTTCAATTTCACTGTCTGGACGAGCCTGCGGCTAACGAAATTTATATACCCATACAATGCCATCAAAGCCACAAGGTGTTTATGGG
CTTAACTCCGGCCCTGATTCCGCCTACTTGGCTGCCTCGCCTCACAGATTTGGATTTGCATGTGCCACGATTTGGGGCATGCGGCTGAGAAGGCGGCTGAGCAGGCGGATAG
AAGACAAATAGCCAGCACGTTGCGGTTCATTTGGCTGTCACAACAAATGCCGAATTGATGCAAATGTAAATGGATTCGATATCCAGGAACGCGTCGTCGGAATCCCAAGCAG
CAGTCAAGTGATGAATACACGGCTTAAAAGATACTGTATTAACTTACAGTAGCTTACCTTCAAAGCAAAAATTGTAGATTTCATTCTCTAGTAACGAGTTCATATCCATACC
TCTTTAATATTGGATACTATTTATTACCCAGTTTATTACTTAGTTTTTATTTCAGAATCAAAAACAAATATAAGCAAGCTTACCCCATTTAAGTACTTGGCAATTGCAAATG
TGCACATACCACTTTCTTTTCCATCCACAAACTGGCTTTCAAACTGCTGTTGGCCAACTCATAAATATGGCAGACAATAAATCTAGATTCTTTTTCAATTGCAGGGTTTGGC
AAACGATTGGCAATAATAGGAGTCATCAAAACGAGAGCAAATGTCATACAGCATGTCGTCCTTATTCCGTGGCGTTAGTGCGACATTTGGAGTTGGAAACTAGGAGCCCACT
TAGACTAACAACGCCCGTGACCATTACCATGACGCTACTGCAGAGGCTTCAGGCCATGTCGGCCACCACGACCAGGACAATACTGGAGGGCAGCATCAGCAGTTTTGGTGGC
GGGACAAATGAGCCTCTGGCGAGCAAAATACCCGTTCTGGAGGAGTCAGCCTCACATGCCAGATATTTGAAATTCATTGCCGACGGGCTCATCGACGAGGGACTGGGCAGTG
CGGTGGGCAGTGGGAGCAGCATCGCCGTATCCGTTGAAGACGTGGTCGCCGGACAGGCGCAGGACATCCAGGCGAGCGAAGGATCCACCGACGACGCCGACGGCAGTAGCCA
TTTGGCATTAGTCTTCGTCAAGTGTTTCATTATTGGTTTCATCATACTGGCCGCCATCCTGGGCAACATGCTGGTGATTGTGTCGGTCATGCGGCACCGGAAATTGCGGTGA
GTCCTTGCCATTGTGCGTTCTAGCCATATCCGTTGGTCAGGACATGTGTGTCCTCAAATGTCTGGGTTTCATTAATTTGTCGCCATGAATCAGGTGGCTTTTTTCCCTTTTT
GCGCCATGGTTTCAATTTCATTGTGTGCATTTGGCAGGAAGCTTCCAACTATAAGTTACTTAGATAATAATTTAACAAAGCATGAACTTCAACATCAAAACATTAGTGAGCA
AATATTAAAAGAGCCCTTACTTTCGCTTTTGTTTAAGGTAAATTGCCCGAGCTGTAACCTGCCACATTGTCCACGAATCGTGGGTGTTTTAGCTGTCAAGGACATGACTTGT
ACACCCCACCCCCCCCGCCCCTTTTTCCTTTTATTTCTTTGTTTGCTCCTTTTTGCCGAACGATTTTGTGCGATTTGACCGCGTTCGGCTTAACGAGTAGCTACCTCCTTTT
TTTGCGGCCATTAATTTGTCATCATTTATCATATTTTCCATCATTTTTGCGTGTCAAGCGCACAGAACTCATCTCGTTCGCCCACCGAACACTTCAGAGGGGGGCGTGGTTG
GTTGACTAAGTTAAATGCTAGATACGTGCGTAATTACATTAGAAATGTTAATAAGGGGCGAGCATAAAAAATATTGTTGCCTACTTACGGGGGCGCCATTGGCATTTACTTTA
TGTGCGGCCATATAATTTGTCAAATTTATGAATTGGGCGCCACCTTCACATGAGATATGGCCACGTGGGCGTTTGAGGGATTTCCAGGTATCCAGGTAGCAAAGTACGTACG
TGTTTCAGTCGCAAATTCAAAAGCCCAACACCGCTGGCATTTAAAGAACCCAACATAAAACACTTGGAGAGAGGCGACAATGGTGTAACCTAACCCCGGGACATTCAATGTC
AAGGCGCAGACTATTAAAGCAGCCCCGGGAAAACGTCTTCCACTGTCATTGCCGTATTCCTTTCGCCTTGGAACCAACCAACGCACACCTGACATTTGTGGCTGCCACCGAG

AATGAGGATGGGCATGGTGATGAGCATGGGGGATGAGGCTCATTTAGCCTGCCAGGGACAGGCAATTTGGGACCTGACCCTGACTACTACGTGTATGCCATGTTGCCGTTTT
TATCTCAACTGCCTTGCGCCAGCATCTCTTTCATTATTGTTGGGCTCATTGTCATGTGCGTTTTCCGCACCATTTTCACCTTTTCCGCCTCGCTAGGCAGTTTAAATCATCA
CGGCAAATTAGACTTAATTACGAAACAACATGCATTCCCACTTAGGTCCCATGCTACCTAAAAACCTTGTGACCAAGCGGAGTAGCAGGTATCTATTTCAATGTTCTGGATA
AGCCATCTGGGCGACAGTTGTGCGACTTTTTCACGACTGGCCACTCAACTTACATTATTCATGTCACACGCATCGATAAAGTGGGTGACAAATTAAAGTCAGACTGCAGATC
AAGTACGAGGAAGAGCCGGGAAAGTGTTTCCTCACTTTTCCTATTAATTAGAGCTTTCCACGTTGGCTGCCTGCCGAAAGTGTCGACTCATAATTGTAATCAAACTTGACAC
CCAACAAGTTTCCTCCTTCATTCCCGAGCACCCACTATATAATCGGCTTTTGGGTAACGCCTCAATAACATTTTCGTGTTGCCTCTGATGCGGGTGTTTTGCCTTTGTACT
TTTTGGGCGCTTTAAGCTTGCTCCAAAAATGTAGACCTTTCGGGGCGTCAGTTTGTGGCCGTAATTAGGTTTGCGGGAGATGGTTTAATGTCCCAGTGAAATATGGCAAAAT
GGCGAGCGATTGTTAACGAGTTTAATAGATAAAGGTCGGATACACATAAAAAGAATATCCATGATTGATGCAGTTTAAGAAGTTCCCATCCAAAGTTCCTTCATTCAAAAGT
TTATCTGGCACATTAATCCGTTTAGACGGAAAGCAAAACTTTCATTTGCGATGTTGTGGTACGAAGAGTCAAGTTCCCGCAGATGTTATGGTTTTAAACATGAATTCCAACC
GCCCATTCGGTTGCGGTTATATCCGTGGAATGGCGTTATGTACGCCTGATTGTTTGGAAATCAGTGCGGACGGCCGATAAACGCATTTGCTGCTCCACATAGCGTATACGCC
GCGTGCACCTAATTGACTTAACGTTTTGCGGTGCCACGCAACGTTGATGATCGCTGGTGGAAGTTGCGGTGGCGTTAAGCTAATTATTAATTAACCCATTAAACAATGAAAT
CCACTCAGAGAAGCGCCTTTTCGTACATGAAACGGCGGTCACACGCGGCGTATGCGCAATGTGCGCACCAAAGTTTCCAACTCAAGTTGGACAGCAATCCGAAGTAACGATT
AGCGCTGCGAGCTCTGAGTGACTTGAAAGCTTAATGCACAAATTTAATTTGAGAACTTGGCACTCTCAATACCACTGGCACTCAGGCTTTAGGTTCATTCTGCCATTTTTGG
TTGACTGTGTGGCTTCCGCCTTGGTTCCTGCTAAGTTTTAACCTAGCATTTACTCCTCCTCCTTGCAGCATCATTACCAACTACTTTGTGGTCTCTCTGGCCGTCGCCGACA
TGCTGGTGGCCCTCTGTGCGATGACATTTAATGCTTCCGTCATGATCTCGGGCAAGTGGATGTTTGGTTCCGTGATGTGCGACATGTGGAACAGCTTCGACGTCTACTTCTC
CACCGCCAGCATCATGCACCTCTGTTGCATATCGGTCGACAGGTGAGTCCCGGGTTTACTTTGTCATTGTGGAATAATTGAAGTTCCGAGTGCTACTGGGATGTCTTTTATA
CCTTCCTTTAAATGCCCATCAAATTGAATAGTTGCTTGACAGTTTCTAGTGACAATCTTTTTTCCCCTTTACCCGTAGATACTACGCCATTGTGCAGCCACTGGACTATCCA
CTAATCATGACACAGCGACGCGTGTTCATCATGCTATTGATGGTGTGGCTATCGCCGGCGCTCCTCTCGTTCCTGCCCATCTGCTCGGGATGGTACACAACGACCGAGAACT
ACAAGTATCTCAAATCGAATCCGCATGTGAGTAGCTTACTTGAAGCCCTTAATTGGGACAATTAAATATTTGTACAATAGTGTATAACTCTATTATGGAGCCCACAGATTTG
ATGAGTAGGTCGATCCGGTGTATGCACCTACATTATTTATTCAATTTTCTTGCACATCAAACTAATAGAAAATTCTAATTTAGTTGAAACGGTATCGACTTTGCATGCATTT
TGCATGCCTTCGATGCCACGTGGGCGTCCATGTGGCAGTGGCAACCTTTGACTTTCCCCCGACCCGCCTCAACGGAAATTGATTTATGTTTGTGTGGCTGAAGCAGTTTTA
CGCCTTTCAGGCTGCCAGCCGACAGGTAGCATATTTATTACTCCATTTTATTGGTTTCCTCCGCAAATACTTGGCAAACTCCTTGGCAGCACCTCCTCCATCTACAAATATA
TACATTTATATATATATATATATATATATGTTCTCTGCCCGTCAGCAGCTGCTGGTACTGATTCCCCCCGCCACTGCGGTTATCATTGCGCAATAAAATGCTGATGTAGAAA
TAAATGTAGGTAATTCCATTCCGGTTTTCTGCGCTCCCATTCAGGCTGACTTCAGCATAGGCTGTGCATAATTTTCGGCTCTACACCTGGAAGGACGCCTGAAACGCCTGAC
AGGCCTTTGCATACATTAGCATCCGCGGTGGGCTTTAATTAGTTTTACTTGTTTATGCTAATAACTGTCAGCGCGAAGCCGGCAGAATGTCCTGGCCGTAAATCTGATTAGA
TGTCTTCTGGTGGCCTTTTGTCATGGCAGCGCTGTAATGCACAGCGTAATTGAGTTTTTATTGTCGCACTTGGCATTGCAGTAATCGCATTGCCATTGCAATTGCCGTCGCA
ATAGCCGTTGCTCATACGCAACGTTGGCCGCCGCCAGCCGCAGTTGGCAATTGCCGAGCATTACGCTCCATTACGGTTCGCCACTGGCATCCGACTGCACTGCTCCAACGTT
CATTTTGATTTATTGATAGTCGCTAATTGAGTTGCCCATCAAATCGCAACGATTTCGATTCCGATTTCGATGCTCCCCTCACTCCATATTTCAGATATGCGAGTTCAAAGTG
AACAAGGCATACGCCATAGTCAGCTCGTCGATGAGCTTCTGGATTCCCGGCATCGTAATGCTGTCGATGTACTACCGCATTTACCAGGAGGCCGACCGACAGGAGCGTCTGG
TGTACAGGTAATAGCTATTCGAGTGGCCCCTCCATTTGCCCGCTTCCAATTAACTCTAATTACTTACTGTGTGCGCCCAGATCCAAGGTGGCCGCTCTGCTGCTGGAGAAGC
ATCTGCAAATTAGCCAAATTCCCAAGCCCCGGCCGAGCATTCAGGTGGAGCAGTCGACCATCTCGACGATGCGGCGTGAGCGGAAGGCCGCCCGCACCCTGGGCATCATCAT
GAGCGCCTTCCTCATCTGCTGGCTGCCGTTCTTCCTCTGGTGCGTATACTTGATGGCTCAAATTGGGCAATTATCGCTAGCACTCTAAGGATCAAGTGCCGAGAAGTTTTTA
ATTAACTCGGGCTCATATCAAATAATTTGCACTTTCCATGCTACAGGTGGGCGGAAAGTGTGTATAAACAATTAACAATACAAGTATGTTTCCTTGGGATTGAAATATAAGA
AAGGCAATGTAGACTTAAGCAGTAGTTATTATGCACTATTCAAGTATCTATTCCACCTACTTAAAAGTAAAGAATCTAACACTTTCCCTTTCTCCAATTAAACCGCAGGTAC
ATCGTATCCTCGCTGTGCGATAGTTGCATCACTCCGCGCCTGCTCGTTGGCATCCTGTTTTGGATCGGCTACTTCAACTCGGCCCTGAACCCCATTATTTATGCATACTTCA
ACCGCGACTTCAGGGCCGCCTTCAAGAAGACCCTCAAGGTGAGTTGCGGGCTGGCCATCGAGTAATTTGCATATACCTTCCACTTCGACATCTTGATGGGACTTTTATTCTA
TTCAATTTCGCTTTGTTCGGGCATCCAAAACTGAATTCTCTGCACCACCTTAGACTTAGTACAGGGCTTCGGCAACTGCTGCAATCGCAAAGGACAGGGCCAACGCCACCCA
TTGCACAGAAGCGGCGGCGGCCACCATTTCGACAACAGTATCGACTGGCAACGCGGAGGCGCCACCACCAGCGCCCCACCGACGAGGCTCACGTTTAGCAATTATTTTGGTA
GCTATTGCATAGCATCACGGGTGCACTGAGAAAAAACATGAACAATTTATCAAATATATTAAGTAACTTGGAGTAACTTAGTTCCAGTTACACGACTAACAAATAAGAAAAA
AGATATGGAAAAAAGAACATTGTCTCCTCATTTTTTCTCAGTGCAGCTCAGTGTTCCGCGGAGAGCGTAGATAGCGTAGCGTTCATATTAGTTGATTGAAATGCTTCTTCAT
CCACATCTACTGGTTGCTAAGCAAACAAAATAGCTTTTCCTCTGCCTGAAGAACTGAAAACTCTGCACGTTCTTTGCTCTACAGCTCTACAGCTCTGCCTGCGCCTGCCTCT
GTCTCTCTATCAGTTTCATTCTCGAGGGTATCTTTCGGGCATGGTGAATGGGACAGTGAATGGGGGTTGCAATGCTTGGGCTTTCTAATCGCTTGAACACTCTGCTTGCCAC
TTACCATCGTATACTACCATTTGCCATATATAGTGCCATATACCATATATCATTTCTCTTTGCACGCATAGCTGGCGCTTACTGTGCTGATTCGTTGCCTATGGCCCAACAA
AAGCTAAGCCCGCCAAAGTATGCAACAGCAAAACAGAAAGCATAAAACGTAATATTTAAATTAAAACGGAACCTCTTCTTTGTATGTATGCCCAGCTACAACCAGCTACACC
TACACCTTCTCCACATTCATACAAGCGTAAATGCTGATGATGAAATGTTATTCTTTCTCTAAAAATTCTCCCTTTTATGCTGCACACTTCGACACACCAGCAACACCAACAC
CAACACACACACACACACTGAGGCACATGGACGGGTCCTTGTTCGTCCCTTTTTTTTATATGAATGCGCGGCCTTTCATTTCGTTTAGCATTTTTTGTGTGAGTCTGTTGTT
TGTTTTTTGTTTCTTTCCCTTTAGCCTTTGTTTTCTGTGTGAAACCAAATTGTCACAGACATGTGCCGCCAGCACCCGAAAAAAAAAAACACCATACCCCTGTCGATTGTGT
CGCCCGTCCCACGCGGCGTATGCTTAACCTTTGAGTGCTGTACAGTAGCGAGCTAACTCTATTAAATTACATAGCGTACTCTAACGGTAGCATGCGAGTTTATCTCATCTAC
TCATAGTTGGGTGTCCTCTTCATTCCACAGAGTCTGTTTCCCTACGCTTTCTACTTCTGTCGACGTGGCAGGGGGCGAGACGATGACCGGGATCTGGAGTTCGGCGGTCCCA
GCCGACGGGGAACCAATGGAGCCCAACGGACCGGATCCGGATCCGCCGAGATGGCCAACTGCGTCAACTCCACGGCCTCGTCGGAGATACACATGAGCGTGATGCGTGCCCG
CCAGTATGCCGTCAATGTCACACCCACCACGGACGCCCAGATGCAGCAGCTGCATCCCCTGTACACCAACTAAACGAAAAACACATATGGCGTATACTTGATTTCAACTCCA
GTGCTTCACAATGTCTGTTCATGAGATGAGTGCGCAGTGCACAGCAAGTGTTCTCCAGAAAACCAGATTGCCAGCTTACAAATACAAATCACTCGATAAAAACATTCCTCAA
AATGGCTGCTAAGAGGGCTAAACAAGAATTTAACCACAAATAATAGCTAAAGCAACAGCCACAGCAACAATACGACAATCCCACAGATTAAATAAAGCCTTCAAGAGCAGCA
AATTGACGTGCAAAATAAATTGAAATCTTATAGAAGTATGTATTTCGACATACAATACATATTTAGGTGTCATGGAGTCAGGAATTTCATAACGAGATGTTTATAAATGCGC
ATGCAAATAGATATATGAATTTGAATGTTTGATATACGTATATAAAAAGATAGATGTTAACCCACAAAAAAAAAAATAATAAAATATGAAATATGCATTGATTGGTAATCTA
AAACGAATCGCAAATTAAAGCTACACTGTGCAACAGTGGTATGTAGTGAGAAAAAAAGTGTAGGTTCAAGGTCTGGTTGAAAGAAAAAAATTTTTAATCTGTTCAATTATTT
GAAATATTAACTTTAAATACGATCAAACTTATCTGGTCCTTTCACATTTTATTCGATTAGCGAATAAAGACTGACAATTTTATTTTATCAGTGACTTAAAAATGATTTAATT
AAAACATTATCCACATATTCAAAGTACTTAACCAGGCCTTAAATTCACTTTAACTCTCCAATAAACAAAGATAACTAATGTTAGTTGTAGCAGAGTGTTCGTATTGTTTGTA
AAACTCATACTCTCAGTTCTGTTAGATGTAATGACAAAGCTTTGATTTGATTCGTTAAATACATTAAAATAATCTAGTTCATTGAAAAACAATCAAATTATATATGTCATAA
TACTTACATACATAGGTATTCATATTCATAAAGAATACATATAGGGCCTAAAGCGCACGACTAAC
(SEQ ID NO: 67)

Exon: 1001..1564
Exon: 4101..4298
Exon: 4447..4618
Exon: 5695..5831
Exon: 5905..6087
Exon: 6381..6534
Exon: 7983..8249
Start ATG: 1149

Transcript No. : CT21432
GGTTTGGCAAACGATTGGCAATAATAGGAGTCATCAAAACGAGAGCAAATGTCATACAGCATGTCGTCCTTATTCCGTGGCGTTAGTGCGACATTTGGAGTTGGAAACTAGG
AGCCCACTTAGACTAACAACGCCCGTGACCATTACCATGACGCTACTGCAGAGGCTTCAGGCCATGTCGGCCACCACGACCAGGACAATACTGGAGGGCAGCATCAGCAGTT
TTGGTGGCGGGACAAATGAGCCTCTGGCGAGCAAAATACCCGTTCTGGAGGAGTCAGCCTCACATGCCAGATATTTGAAATTCATTGCCGACGGGCTCATCGACGAGGGACT
GGGCAGTGCGGTGGGCAGTGGAGCAGCATCGCCGTATCCGTTGAAGACGTGGTCGCCGGACAGGCGCAGGACATCCAGGCGAGCGAAGGATCCACCGACGACGCCGACGGC

AGTAGCCATTTGGCATTAGTCTTCGTCAAGTGTTTCATTATTGGTTTCATCATACTGGCCGCCATCCTGGGCAACATGCTGGTGATTGTGTCGGTCATGCGGCACCGGAAAT
TGCGCATCATTACCAACTACTTTGTGGTCTCTCTGGCCGTCGCCGACATGCTGGTGGCCCTCTGTGCGATGACATTTAATGCTTCCGTCATGATCTCGGGCAAGTGGATGTT
TGGTTCCGTGATGTGCGACATGTGGAACAGCTTCGACGTCTACTTCTCCACCGCCAGCATCATGCACCTCTGTTGCATATCGGTCGACAGATACTACGCCATTGTGCAGCCA
CTGGACTATCCACTAATCATGACACAGCGACGCGTGTTCATCATGCTATTGATGGTGTGGCTATCGCCGGCGCTCCTCTCGTTCCTGCCCATCTGCTCGGGATGGTACACAA
CGACCGAGAACTACAAGTATCTCAAATCGAATCCGCATATATGCGAGTTCAAAGTGAACAAGGCATACGCCATAGTCAGCTCGTCGATGAGCTTCTGGATTCCCGGCATCGT
AATGCTGTCGATGTACTACCGCATTTACCAGGAGGCCGACCGACAGGAGCGTCTGGTGTACAGATCCAAGGTGGCCGCTCTGCTGCTGGAGAAGCATCTGCAAATTAGCCAA
ATTCCCAAGCCCCGGCCGAGCATTCAGGTGGAGCAGTCGACCATCTCGACGATGCGGCGTGAGCGGAAGGCCGCCCGCACCCTGGGCATCATCATGAGCGCCTTCCTCATCT
GCTGGCTGCCGTTCTTCCTCTGGTACATCGTATCCTCGCTGTGCGATAGTTGCATCACTCCGCGCCTGCTCGTTGGCATCCTGTTTTGGATCGGCTACTTCAACTCGGCCCT
GAACCCCATTATTTATGCATACTTCAACCGCGACTTCAGGGCCGCCTTCAAGAAGACCCTCAAGAGTCTGTTTCCCTACGCTTTCTACTTCTGTCGACGTGGCAGGGGGCGA
GACGATGACCGGGATCTGGAGTTCGGCGGTCCCAGCCGACGGGGAACCAATGGAGCCCAACGGACCGGATCCGGATCCGCCGAGATGGCCAACTGCGTCAACTCCACGGCCT
CGTCGGAGATACACATGAGCGTGATGCGTGCCCGCCAGTATGCCGTCAATGTCACACCCACCACGGACGCCCAGATGCAGCAGCTGCATCCCCTGTACACCAACTAA
(SEQ ID NO: 68)

Start ATG: 149

MTLLQRLQAMSATTTRTILEGSISSFGGGTNEPLASKIPVLEESASHARYLKFIADGLIDEGLGSAVGSGSSIAVSVEDVVAGQAQDIQASEGSTDDADGSSHLALVFVKCF
IIGFIILAAILGNMLVIVSVMRHRKLRIITNYFVVSLAVADMLVALCAMTFNASVMISGKWMFGSVMCDMWNSFDVYFSTASIMHLCCISVDRYYAIVQPLDYPLIMTQRRV
FIMLLMVWLSPALLSFLPICSGWYTTTENYKYLKSNPHICEFKVNKAYAIVSSSMSFWIPGIVMLSMYYRIYQEADRQERLVYRSKVAALLLEKHLQISQIPKPRPSIQVEQ
STISTMRRERKAARTLGIIMSAFLICWLPFFLWYIVSSLCDSCITPRLLVGILFWIGYFNSALNPIIYAYFNRDFRAAFKKTLKSLFPYAFYFCRRGRGRDDDRDLEFGGPS
RRGTNGAQRTGSGSAEMANCVNSTASSEIHMSVMRARQYAVNVTPTTDAQMQQLHPLYTN*
(SEQ ID NO: 69)

Name: 5-HT1A/Ocr-like receptor
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384515
CTGGACGCCGGACACGAAAGCAGATCACCTGCCGTACAGTGCCAAGATCATCGAACGCATTTACAATGAGGAGATCGGTGATGGCGGTGGCCACAGTGCACGGAGGATAAAC
ATGCTGGAGTTCAGTCAGTATCTGGAGCAGTACCTGTGGCCACACTACCAACGTGAAACCGCCACCCATGCCCACCTTATGTCCATCGTTATAATGGCCAACGAGAAGTTCC
GGGAGCGCGTCGAAGTTTGGACCGTGTTTGAAAAGCTCCCGGATCAGTATCCAGCCTTCTTTCGCCACGTGCTAGAGAGCTGTTTGCCGAGCAAAAAGGCAAAGGAGGCTAG
TAGCACGCTAAGGGAGCGGACAGCGCTGTTGATGTTCATTAACCACTGCTTTAACAGCATGGAGATAGAGCTGTGCAGAGAACAGGCCAAACGACTAGTCTCCTTGTCCATG
TGGCATTGCCTGCAGCCTCGTAAGTTCTTCTTCGCAACTCTTCACATCTATCAAAGTGATCAAAAATGTATTTGCTTGCAGGCCGTCGCGAACAAGAGCTTCGAGAAGTTCC
TGAGTGGCGAAAGTACTGGAAACGCTTGCTCAAGAAGGAGAAAGACAGCAAACCGGAGGTCCTCTGGGAGAGACACTTTATGCAGAACCTTATCATTGACTTCCTGCACATA
CTCGAAAGCATTCCCGCCGAAGGCGAGGTGCCGCGCAACGTGGTTCACTACTGTGAGCGCTTCCTAGAGTTTATTATCGACTTGGAGGCACTTTTGCCGACTCGACGCTTCT
TCAACACGGTTTTAGACGACTGTCATCTGATTGTGCGGGCTCTTCTGTCACCCTTGGTTCGTCGCGAGGAGGGAAAACTGTTTGGTCAGGTAGGTGTCAGACTTCTCATTTA
ATAACTATCTCGTACAGCATAAATTAGTTAAAATTAAGTACAAATCTCAATGTTGGACATTGCTAAACTACGAAAACACTTATATTAGTGTGGACGGTGGCTCCTCAGTAAT
CGTTGCCGTTCATATAATGCAGCTCATCGCCCCAGTCGTTTGCCGACGGTGGCTCATTGTAGCAGCAGGTCTTTTTCAGCAGAAATGTCACATGTCGCTGGCGCAGCACGCA
TATGTATAGCAACAGCGGCGTCTGCAGGGCGTTAACTACGATGTGGGCGTAAAGCAGGCCTTCCATCTGCAACCAGGACATTATGAGGAATAGCCAGGCTATAGACATTACT
AACAGCATCAGCGAGAACATGATAAAACTGCAAGGAAGATAATGTTTTCATATACTTATATGTAATGCCTAGATGCCGATCTACTTACTTGGCTTTCAATTTGTGCGCAATG
CGGCCGTATACCGTCCTGCGGTTAATCAACTTCCTGGTGGTCACATAGAAGAATATGTTTACCAAAATGGTACAAGCAATTGGAGCAAAGAATATGCATATGCCTAGCCAGC
CGATCGTCTCTTGCTCGCCCACCATGTGCTCCTGTTTGTAGGATTCGGCGTCGAGGAAGAAATGTGCGAAGACAGCCAGTGCAGCCATCGTTGCTGTGCATCCCCAAGCATA
CGCAGAATAATAGCAGTACTTCCTGCCGTCCGTGACTCGCAGGAAGACGTTCCTCGAACGAAAGGTCTTCCAAATGTAAAAACCGAAACTGTTCAGCCAGAAGAAAGCTGCC
AAGAGGCTGAAGCACAGGATAATGTCCGCCACAATGAAGCTCACATGGCTGGTCAGCTCGGTGAAAATTCTGACAAGGTCTGCTGCCTGGCTTACCATAAGGCACATGGCTA
TCGTTGTCACAATGTTACCAACCAGATCTCTGCTGAATGGATTGAAAATTAACGCATGAAGCAAGTCAAATTCAAACTATTGCTCACCTGAGAGTAGGAAGTATAAAATAGA
TGATGGCGATGACCAGCAGGATAACAAGCGAGATGCCATGGAAAATTGGGTTTAAGATTTTGCGGAGCAGAAAGTTTGAGTCACTCCACTTGATCTCCTTGCGGGCCAAGCA
AATGTTGGCGAATAGCACATTTTCCTCTCCAGTACTGGAAGTGGCCTGCGAATGGATTTTATGATCATCCATGCAATTTATGTGGACTTTAGCTACTCTTACCTTGTCAATG
CAATACAATCCCTTATCGTAGTCATGGTATAGCGGTTCCAGACTGCCGGCAATATCCTCCTCATAATCCGATTGTATTCCATGACGCTCCTCTGCTTCGTTTTCGGCATTGG
TGTAATGCCTTAGCCTCCCATCATCCAGTAACACTAGCTTATCAGAGCTCTAGAAAATATAAAAATATCTATAAGTATCTAAAAGTTTCTATGCAAAGAACTTCAAACTCAC
GCCAGCATAGTGGTATATTGGCCACATCTGCATGGATCCACAATTTGGAATCCCGATGACGAATTTGAACTTAACGGGCCTGTTTTGCTCTCCTCCATAGCTGGTGAACATC
GGCTGGAAGTAGTCTGCAAATGGAAACGGACGATAAATAGAGAACACAGACCGTGGAAAATCATAAAGGCAGTTCAAGTTAACGCTGCTTTTGTTGCGGCTGGTGGATCAGA
GGGGAGCCAGATTTTCTTTTCTCTCTGCGAATATAAATATTTACGCTTCTGGGCAAAGAACTTGAGCTGTTCCATGCGCCTTCATCAGCTGGTTTTGCAATTTGTGCAGATT
GGTTCTTCTTCAATTATCGGAGTTCGGATACATTTTTTTCCACTCAGCCATTCATATGTTAATCATAGCTAAGCACACGGAAATTAGATTGCATGTTGCGATTATGTGCCCT
GTTGTATTGTAGTATTGGGATTACTCAAATGACTACAATTAAAATCGGATGAGATCAGTACAATGCAATGATTCAACTATCAGTAATAGTTACATTTTTATGTGTCACCTAT
ATGGCCGATCCATGGTCATAAATGTTTTATCCAGTTTAGTTGTCCGAGCTAGAATATCGATATAGATGCTAGAGTACGCTTTTCTATTTCCCACAATTGTTGCGATTGTTT
GTCAAACATTCACCCAAGGCAATATATCACTATATTCACCATTAGCTCCACAAACATTAATTACCAGGGCTCGCATTTATGGGGAACAGAAGCATCTTATATTACCTTACTA
ATGGATTTTTTGCCTCATTTACTTCCAAGAGATAATATTCCCGATTTGGAATCCCTCATCAAGTGGCTACTTTATAGTTTAAACAGAAACATGCCCCACAGCATAAATCA
TATTTCTCATCAGGATTGGGGCCACCCTCTGTTTGTTTTAAGAACTTTCTTGGGGGAGTGCGAACTGTGGAGTTGGGGATATTTAGCTGTTGGCTTCGTTCCACTGATCCAC
TGATAGCCAACAAATGGGCCATTTAATGTTGGGCATACATATGTTGCATGCCAGAAGGGAAAATCCATAAAGTTGAATGGCCAACCAACTTCGTTGTGGTTGTTATGGCAAC
TTTTTCGAGCGACTAGGAGTTTTCCGCCTGGCCAGCTGTTGAACTTACCCGTTTCATTGACCTGCTGACACTCATGGTCCACGTGTATCTCGAACTTCTCGCAACATTTGTT
GACCAGCACTAGGTTGGGATCGGAAGATAGGGCTGTTGATGATCCGGCGCTAGTGACATGACTGGTT
(SEQ ID NO: 70)

Exon: 2651..2621
Exon: 2477..2353
Exon: 2289..2119
Exon: 2061..1880
Exon: 1824..1321
Exon: 1259..1001
Start ATG: 2651 (Reverse strand: CAT)

Transcript No. : CT21585
ATGGAACAGCTCAAGTTCTTTGCCCAGAAGCACTACTTCCAGCCGATGTTCACCAGCTATGGAGGAGAGCAAAACAGGCCCCGTTAAGTTCAAATTCGTCATCGGGATTCCAA
ATTGTGGATCCATGCAGATGTGGCCAATATACCACTATGCTGGCAGCTCTGATAAGCTAGTGTTACTGGATGATGGGAGGCTAAGGCATTACACCAATGCCGAAAACGAAGC
AGAGGAGCGTCATGGAATACAATCGGATTATGAGGAGGATATTGCCGGCAGTCTGGAACCGCTATACCATGACTACGATAAGGGATTGTATTGCATTGACAAGGCCACTTCC
AGTACTGGAGAGGAAAATGTGCTATTCGCCAACATTTGCTTGGCCCGCAAGGAGATCAAGTGGAGTGACTCAAACTTTCTGCTCCGCAAAATCTTAAACCCAATTTTCCATG
GCATCTCGCTTGTTATCCTGCTGGTCATCGCCATCATCTATTTTATACTTCCTACTCTCAGCAGAGATCTGGTTGGTAACATTGTGACAACGATAGCCATGTGCCTTATGGT

AAGCCAGGCAGCAGACCTTGTCAGAATTTTCACCGAGCTGACCAGCCATGTGAGCTTCATTGTGGCGGACATTATCCTGTGCTTCAGCCTCTTGGCAGCTTTCTTCTGGCTG
AACAGTTTCGGTTTTTACATTTGGAAGACCTTTCGTTCGAGGAACGTCTTCCTGCGAGTCACGGACGGCAGGAAGTACTGCTATTATTCTGCGTATGCTTGGGGATGCACAG
CAACGATGGCTGCACTGGCTGTCTTCGCACATTTCTTCCTCGACGCCGAATCCTACAAACAGGAGCACATGGTGGGCGAGCAAGAGACGATCGGCTGGCTAGGCATATGCAT
ATTCTTTGCTCCAATTGCTTGTACCATTTTGGTAAACATATTCTTCTATGTGACCACCAGGAAGTTGATTAACCGCAGGACGGTATACGGCCGCATTGCGCACAAATTGAAA
GCCAATTTTATCATGTTCTCGCTGATGCTGTTAGTAATGTCTATAGCCTGGCTATTCCTCATAATGTCCTGGTTGCAGATGGAAGGCCTGCTTTACGCCCACATCGTAGTTA
ACGCCCTGCAGACGCCGCTGTTGCTATACATATGCGTGCTGCGCCAGCGACATGTGACATTTCTGCTGAAAAAGACCTGCTGCTACAATGAGCCACCGTCGGCAAACGACTG
GGGCGATGAGCTGCATTATATGAACGGCAACGATTACTGA
(SEQ ID NO: 71)

Start ATG: 1 (Reverse strand: CAT)

MEQLKFFAQKHYFQPMFTSYGGEQNRPVKFKFVIGIPNCGSMQMWPIYHYAGSSDKLVLLDDGRLRHYTNAENEAEERHGIQSDYEEDIAGSLEPLYHDYDKGLYCIDKATS
STGEENVLFANICLARKEIKWSDSNFLLRKILNPIFHGISLVILLVIAIIYFILPTLSRDLVGNIVTTIAMCLMVSQAADLVRIFTELTSHVSFIVADIILCFSLLAAFFWL
NSFGFYIWKTFRSRNVFLRVTDGRKYCYYSAYAWGCTATMAALAVFAHFFLDAESYKQEHMVGEQETIGWLGICIFFAPIACTILVNIFFYVTTRKLINRRTVYGRIAHKLK
ANFIMFSLMLLVMSIAWLFLIMSWLQMEGLLYAHIVVNALQTPLLLYICVLRQRHVTFLLKKTCCYNEPPSANDWGDELHYMNGNDY*
(SEQ ID NO: 72)

Name: mth-like 5
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013383774
GCTGTTCGCACGGCATTTTTGGGCCAGCAAACAGGATGCGGAAAACCGGAAGCTGGGAACTGGAAATGGGCAAGAATCTGGTTTTTGGCTGCCATTCAATTAGGTATCGGTC
AGCCCAATTCAACTGCTGCTAACTGGCAACTGGAGATTTGGCGATTGCAGCTGCCCATTGTCAACACTTCAATTAACGCAATAGCAGGTAAAAAGGGGACATCTTTTCATCG
AGATCTCGACCTTGTTTTGCTATCCTTGCAAACAATTGATGGTATGTCGTAAGTTTGTGCGAATAAAATGGTTTTGGAAATATATATATATATTGTTTATTAGTTTATGGCC
CCAAAAAATGCTACAACACAATATTGAAACCGTTGTTATCTGAAGGGAATGCAACGATTCGAGCTTCATATTGAGAAGGTCCTGTTTTCTCTCTGTTCCTGCATAGTTCCTC
TTTCATTTCGCATCTTTTTCCACTGGCCGCAGTCTCGCAGGATATGCTAACTAAGGACTATGCCACCATATTCGACAATTTGTTGGGCACCAAGTTAATATGCATTGATGCG
TTATTCATACGCCCCGTGGGACGGGCGAATAGCAAAACTCTCAATGGGCCTTTGTTAATTTGTATGTGTGTGTATGTGTATTTTGCGTGTATTTGATTTGATGCGTTTAGTA
CTTAGGATAACAATTACGTTTAGGTTAACTAGCATAAATCTATAATCATCACATTATGGCACAATGTATGTTATGTAGTATATATATATCGTATATATAGAGAGACTGATAT
ACATAGATATCGATTGATATAACAAATTAAATTTTCAGCTAACCGCACGCTTCAACTGAATCACTTTTGATTCAATCATCTCACTTGTGATTCCGCGTGATTTGCTGGCAAA
TTTAAAGCCCTCTTCTCGTGTTTCATGAGAAGTTTCATGATATCATCTATCACTCCTCTCATGGAGTTTAAAACCATAAATATTTTATTATAAATATTCACTATTCACGCTA
TCAGGCGACCCGTATTACGATTTCGACGACTTCCGGCTCCGCCATAATTACTACCATAATGATCGCCATTTATTTGGGTCAGTGTAGTCGTTGTATGCGGCACATAGAGACT
CGTATTGTTACGCGCATGTCGACGTCTATAGCGATATCCCGTTATCAGTTCTCGAACCGTCTTTCGGTACTTATCGGAAAGTACACAGTATAGAAAAAAATTGGATGCCGCA
TGTAGGGAGGCCAATAGGTTACACACATTTCCCGCCACCACCTGAGTGCCAATTTGAATTTCCATATACTGATTGACTATCAAATAGATCGCCCAAGGCAATTGGCAAACTA
TGAACATTAACACCACCGCAATAAGTGTGATCGTCACGCGATTTTCCTGCGATACGCTGCCTTTTAGTCCAGAATTTGATTTGCGCTAAGGATTCAGAATTTATACAACATA
TCATCATCATCATCATCGTTTCGAATAACTCACTTTTGTGCGCCTTATGCTGCTGGCATTGGTCAGGTCACCACGCAGATTCTTGGACCGATGAACCAATAGAATGAGAATG
GAGTTGAAGGTGGCCAATAGAAGTAGGGGGAATACCACGAAGACCAGCAGGGTGAACAGGTTCCACAGATAGCTGTATAGCTCATCATAGCCCAGTTCCGAAAGACCATGAT
TGTAATAGGTCACCTAAAACATATAGATAGATACATAGAATACATACATAGATTTCCATTAAGAACAACCTTTTAAGCAATGGAATATCCCAAGCGACTAAGTAAAAGCATG
TCTTACAAAACCCAAAGCACTGATATCATTTTAAGTTATCACCGAATTTATCATAGTCATGGCGCTTGTACACCGCTCGTGATTTTCACTGGTTAACAATTTTCTCATGTTT
TTTTCTCCAGGGAAATCGACGGTATTTCCAACATTATTTTACTCACATTTTGACAGTATTCCGTTACGTTAAATGCGAATGCATCCGTATTGTTGATGTTGTGATTCTCAGC
GCTGCGCTTGCGACGCCACAATTGCAGCAAACTCTCTGTGACCCCTGACCTCTGGGCCGCGTCTGCCACCTTCTGATCCTGCTCCTCCAGCGTTTTTCTCAGTGTGGGCAGC
GTCACAAATCCAGATGACTGCCAGTGACGGCGATGTCGTGGAATGTTCTCTAGTTGAGGGCAAAGAAAGTAATGCAACTATCAACTACTTCACTTATAAAGCTGTCTAAAAG
TATGCTTCGTACTTTTCAATGCATACTTCTAAGCATCTCTATTAATACAACTTCGATTTTTCCAAGTGTACACTACTGCTATACCCTCGTTACTCACCTGGCTCTCCATCAC
CACTGCCACTGCCCCAGTCGACGAGCTGATTAGAGCTGCCGGACGTCGTGGACTCCATGGCGGAGTCCGGCGATTGCCAAATGGTCGCCGGCGTGGGCAGCGGGGGCGTGGC
CAGTGGGGGCGTGACTGCCACGGGTGGCGGCGGCATGGGCGAGATGTTGGCCACCGTTTGGTTGCACGGCTGGTAGGCGTCATCAATCTGCCTAGTACCAATCGTAATTGTA
TGCTCGAATGCCGTCGATAGTGTGGACAGCAAACAGAAGATGGCCACCGCTGGCAATGTAGAATATAAAGTTGCATATGGAAATGGATTCTATTGTTACATACGCCTATTTC
AGCCAACTCACCTGCTATCACCTTTTTGGCCAGCGATTCCGTGCAGAATGTTTGCCTCTTGAGCGGATAGCGTATCGCAATGAATCTCTCGATGGTGAAGCACACGGCTATG
TAAATTGAGATGTAACCTGTCCGGTAAAAACAGAGAAAGCAATTGTGAATGGGGGCGTGGCAGTACAATACGAACGCAATCAAAGATGCACTGCAAATATTTTGGGAAAATC
CAATCCAAAAAGGATATATATACATATATATTTATTCAGGTTAAGTGTCTAATATAGATTGGTAAATTAAATTCAAACTTCGGTTAGACTGTGCCTTTTTTCCGTGTACTCT
GGAATTATAAATAAACGCAGCTCCGCTGTTAATACCCCAAACCATCCATTTTTGACCGCACCACCAATTTATTTGCCGCGCGGGCTGTCAATTTATTCAAAACGCAGCGCAC
TCTAAGCGATTGACGTCCGTGCTGTGCTGATGAGGTTGCTGCGGGAGGGTGGGATATTCAAGGGGGGTGAGGGGGGTGAAAAGTGGGGTGGGAAAGCGGAAAACGGTTCGGA
AAAACGCTATCGCATGGCAGTGAGCCTCAGCTGACCGCTGGCATTTCGCAGAAAGCGCTGCTTACTATACCTCGTAAATTGGCAGAGCTCAATGAAATTGATGCACAACAGA
AACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCCGCAGGTGCAATTCGAATTGATGGCCCATCGCTGGAATCCGGTTGATACGGGGGAAATGGGGGAT
ACATTGGATCCGGGGATCCAGACGAGGCTAACCCGGGGAACTGGGGAATTTATGGAATCCGCACCATTCCACTTAGTGCTCTCCAAACGTTGCATTCGAACTGTTTTTGTTT
TTTGCCCCCGGCAGCAATTGTTTTCACCGCATGATATGTACGCTTTTCTTATTGTAGTTTTGCGGGAATAATGTTAAATATAAAAACACCCTTTTGGTAAATTTTTTCAAAT
TGAAAGCTTATTTAAGCATAAAATACTCGTGTTTCGATTTAAAGAACACTAATCGAGTTACTTTCGGTCTCACAACCTTATTTATTGCATACTTTGCGACTGCTTTGAATTA
TTGAATTATTACTCCATTGCTACTTAATTCTAATTTGGTCAGTACTCACCGAAACTGTCGCACAGCCAGACAAAATAGCCATAGAGCTGCCAGTAGAGCTTGAAATGATACT
TGGGATAGTCATAGTGCTGGAGTGACAGGATTAACTGGCACGTCAGGTAGGCAATATCCGTAATGGCCAGTGCCGTTAGATATATGTTCGTCGTGCATCTCATGCGCTTTCT
AAAAATACCGTCAAGAATATCGAAAAGAATTGCAAATGGTGTTGATTCTGTTGGGCACTCGACGCATTTGGAAATCAGGCCAAACCATTTCGGTGGACAATGCAAAAACGCT
GCCAAGTGCAACACCAATTGAAATAGCAAATTGCCTGAAAGCCAAACGATCGGTGGGGTTGTGCGCAGAGGGGGTGGAAAATTCCAGGCGGTTGATGTGATGGGCGCGGGGC
CTTGGGCATTTTCCGAAATGCGCACAAATGTCGGCACAGTCATAAAGCCAAAAGATGAACAATAAAAACTAACTGTTAGCATTTAGTTTTCTCCCTGGCATCGCCCGCAGCT
CTGATAGTTAAATTCTCTTGTATTCCCGTAAAACAAAATAAAATAACGTAATGAAAATATTTAATTAAAAAGCCATTCATGCCGTCAACGCTCAATGTTAAGCTAGTTGCTT
TGAAAAGTTGTAAATATGGTTTAAAATAATAAAGTTCATGCTATAAATATCAATAAATTTTGTTTAGCAAATAAATGAATTTGCTAAGATCTTAGCAAGGAACACAGTTTAT
CTTTTTTTTTCACGTTTTCCTGGGAAATGGACATATGGTGTCATAGCAGTCAGCCTAACGCACAGAGCACTTGCAACTGTCGCATATTTTCCGGCAGACATAAAAAACACAC
CCATCCACACAGCATATCGCCTATCTGTGTGCGTTTTTGCATTTGGCACATAAAATGGCTCTTATTGTTTCAGTTACTGTTATTGTTGTTGCACTTGTACCCAACGTAGTTG
TTGCTGTTGTTATTGTTATCGTTGTTGGCTGTGTGGAAATGATTTTAAAGTGCAGCGGTGGCTACAATTAAATGTCGTGTTTAATGCGCTGCGTACCTGCTCGCCGCTTCAT
CCATGTTCCACTTATTTCACACACACAGACACACACACACACGCACAACAGGACATGCAGTTGGATGCACAGAGAGAAAATCGTGCAAGTTAGCTGGCATTTTAAATGGAAA
ATTGTAACTTAAAACTAGTTTAAAACAAGCTCAAACGAATAAGTTGAAGTTTGGAATAAAATTAATTTGAAAAGGCTCATATATTTTATATGCATTGCATACTTTTTAGTAC
TTAACCAAATGCATGAAAATCCAATTAGTCGCTTTTGCATACCTAATAACTTAACATTCTTAACGTGGCAGCTAAAACTTCTTTCAAATGCCATCATCAGCTAAATGATTCA
TATTTAATTCGAAAATCTTGTGACCCTGCTGCATG
(SEQ ID NO: 73)

Exon: 4299..3858
Exon: 2816..2700
Exon: 2625..2338
Exon: 2177..1951

FIGURE SHEET 36

Exon: 1693..1490
Exon: 1189..1001
Start ATG: 4299 (Reverse strand: CAT)

Transcript No. : CT21642
ATGACTGTGCCGACATTTGTGCGCATTTCGGAAAATGCCCAAGGCCCCGCGCCCATCACATCAACCGCCTGGAATTTTCCACCCCCTCTGCGCACAACCCCACCGATCGTTT
GGCTTTCAGGCAATTTGCTATTTCAATTGGTGTTGCACTTGGCAGCGTTTTTGCATTGTCCACCGAAATGGTTTGGCCTGATTTCCAAATGCGTCGAGTGCCCAACAGAATC
AACACCATTTGCAATTCTTTTCGATATTCTTGACGGTATTTTTAGAAAGCGCATGAGATGCACGACGAACATATATCTAACGGCACTGGCCATTACGGATATTGCCTACCTG
ACGTGCCAGTTAATCCTGTCACTCCAGCACTATGACTATCCCAAGTATCATTTCAAGCTCTACTGGCAGCTCTATGGCTATTTTGTCTGGCTGTGCGACAGTTTCGGTTACA
TCTCAATTTACATAGCCGTGTGCTTCACCATCGAGAGATTCATTGCGATACGCTATCCGCTCAAGAGGCAAACATTCTGCACGGAATCGCTGGCCAAAAAGGTGATAGCAGC
GGTGGCCATCTTCTGTTTGCTGTCCACACTATCGACGGCATTCGAGCATACAATTACGATTGGTACTAGGCAGATTGATGACGCCTACCAGCCGTGCAACCAAACGGTGGCC
AACATCTCGCCCATGCCGCCGCCACCCGTGGCAGTCACGCCCCCACTGGCCACGCCCCCGCTGCCCACGCCGGCGACCATTTGGCAATCGCCGGACTCCGCCATGGAGTCCA
CGACGTCCGGCAGCTCTAATCAGCTCGTCGACTGGGGCAGTGGCAGTGGTGATGGAGAGCCAGAGAACATTCCACGACATCGCCGTCACTGGCAGTCATCTGGATTTGTGAC
GCTGCCCACACTGAGAAAAACGCTGGAGGAGCAGGATCAGAAGGTGGCAGACGCGGCCCAGAGGTCAGGGGTCACAGAGAGTTTGCTGCAATTGTGGCGTCGCAAGCGCAGC
GCTGAGAATCACAACATCAACAATACGGATGCATTCGCATTTAACGTAACGGAATACTGTCAAAATGTGACCTATTACAATCATGGTCTTTCGGAACTGGGCTATGATGAGC
TATACAGCTATCTGTGGAACCTGTTCACCCTGCTGGTCTTCGTGGTATTCCCCCTACTTCTATTGGCCACCTTCAACTCCATTCTCATTCTATTGGTTCATCGGTCCAAGAA
TCTGCGTGGTGACCTGACCAATGCCAGCAGCATAAGGCGCACAAAATACCGAAAGACGGTTCGAGAACTGATAACGGGATATCGCTATAGACGTCGACATGCGCGTAACAAT
ACGAGTCTCTATGTGCCGCATACAACGACTACACTGACCCAAATAAATGGCGATCATTATGGTAGTAATTATGGCGGAGCCGGAAGTCGTCGAAATCGTAATACGGGTCGCC
TGATAGCGTGA
(SEQ ID NO: 74)

Start ATG: 1 (Reverse strand: CAT)

MTVPTFVRISENAQGPAPITSTAWNFPPPLRTTPPIVWLSGNLLFQLVLHLAAFLHCPPKWFGLISKCVECPTESTPFAILFDILDGIFRKRMRCTTNIYLTALAITDIAYL
TCQLILSLQHYDYPKYHFKLYWQLYGYFVWLCDSFGYISIYIAVCFTIERFIAIRYPLKRQTFCTESLAKKVIAAVAIFCLLSTLSTAFEHTITIGTRQIDDAYQPCNQTVA
NISPMPPPPVAVTPPLATPPLPTPATIWQSPDSAMESTTSGSSNQLVDWGSGSGDGEPENIPRHRRHWQSSGFVTLPTLRKTLEEQDQKVADAAQRSGVTESLLQLWRRKRS
AENHNINNTDAFAFNVTEYCQNVTYYNHGLSELGYDELYSYLWNLFTLLVFVVFPLLLLATFNSILILLVHRSKNLRGDLTNASSIRRTKYRKTVRELITGYRYRRRHARNN
TSLYVPHTTTTLTQINGDHYGSNYGGAGSRRNRNTGRLIA*
(SEQ ID NO: 75)

Name: TRH receptor like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384839
ATAAATGGCCGAATTTGCAAGTCAGTTCCACTTGCTCTTTAAAAAGCGAATCTCACCGAAATTTAAGCAAAATTTATCACCGACCAATGTGAAATTCCACTAATAATTTTTC
AGCGGTGTTGATGGCTTTATTGACATCGCCAGAAATGATTCGAGCTGTAGCAGCGAAAATTTAATCAAAGACCTAACCCACGAACCTGGACGTTCTCACATAGAAAATGAAA
ACGAAAGCTGAAAGACAGATGGATCGTTCGATGTGAGCTCTTTTTACACTTGGCTTACACGACGACTGCGGCGTCGGTAAAATTAATATACCAAACAATTACACGGCCCGGC
AAAACAATTTTAAGCCCAGTTAATGCCGTCGCACAGTGGCTGCAAGCCACGCAAATTCGCCGCGAAAAATTGCCCACAATTGCCAGTGGCCGAGAAATGAAAGCGCTCAAAT
TTGCATACTCGTACTGTCAAATGAAGCGTAGCAAAAGGAAAATTGCTACACATACTAATTGGAAAAGAAATGCAAAACCAATGCAAATGAAAATAATGACAGGGCGCGGTTT
GCGAAGCTGTCTAATTATTTATTTGACATTGGCAATGACAATGAAACCGAGAGTCCCACAGCAGAGTAATGGATTTTTCTTTCAACTGCTTAAGTTCGTTTTCCATTGAAAT
CTTATTTGGACATATAAACCATACGTAGAATGATAAGAATAATAAAAAACAAGTATGTTGGACACAATTGATAAGGAATCTTTAGTACTTTCTCACCTAAAAAAGGTGATAT
ATATTTGAATTCATGTTGTCATGATAGCTACTCGAACCACTGTGCGTTGGCTACAGTCTGCTGATGCTCACAGTGTTTGAAAGTTAAGACTTTTGCTTTCCTTCCCCCGACT
TGGTTTTTGCTTTTGGAAAGACCAAGCAAATGGATAAACAAGAGTTCAAGAGGTTGTCAGCCCCGCCCCTTTGCGCACATGAGGCAAGAAAGAAAAGAAAGTTATTACAAAA
TTAACAGAGGCTGACAGCTGACAAGGGCAGAGGAATGGGGCCGAAATGGAACTGTATCTAGCTTGAGTTGGGTCCGACCGCCTCGGGCAGAAAATGTGACAAATTTGGAGGG
GAAAAGCTCGTATGAGCGGGACTTACCACCTCCATTGTTGGCACTGGAGCAGTGTTTGGAGCTGAGTCCCTGGCCCGTGTTGGATGCAGAGTTGGGCAGTCCGCAGCCGCTC
ATCACGCCGTACTTGATGGGCTTGAAGTGGAAGAACGAGGGCGAGTAGCCGGAACCGGCGCGACTCAGGGAACTGAACTGCTCCTCGCGGTCCTCCACGTACAGGCCCTCCT
TGCCCATGGCCGCCAGTTGCCTGCCCTTGGGCATGAACATCACCAAGAAGACCGTGGCGGATGTGGCTACCAGACCGAAGGCCACGCAAGCGTCCTTGTGCCGCTCGGCCAC
GGCCAGGCCGCATAGCATCCATCCCAGCCAGATGGGGATGGCTCCTCCGATGGCCAGCCCAATGTAGGTGGCCTCACGGTAGTTGTCCCTGATGCCGCGCGACTTGATGGCT
AGCACGGCGATGAAGACGATCAGAAAGACAATGTAGATCAGCGAGAAGAGCAGCTCCGAGAACTGCGTCTTGCAGAGCGGTATTAGTACGGTGCTAACCGCCGCTATCCGGG
TGTAGATCTCCGGAGTCCCGTCCAGCGTTGTGTAGGACGTGGGGTAAAACAGTGCCGAGTAGTTGGTCTGTGAGGCCACCGTGGTGCTAAGGAAGCCACTGCCCATCACAGG
CACACTGGTGGTGTATACCTCCGGTGGTTGGGTGAGCAGCCACTGCCCTCCGATCGCTACCTGGATGAGCAGCGCAAAGAGGAGCAGCAATCCCTGGTATGGAGCCGGCAGA
TACACCCCTCCATTTAGACTAATCAGAAATACGCATTTGACCAGCAGGGCGGCAAAGACCAGGGCGTAGGCCACGCCCACTCCGAAGCGGATGGCCCCGCAACTGATCAACG
AGGGCTGTGCGGTGATGATGGCCCCCAAACTAGCGCAGGCGAACAGACCGAGAAGCAGCATCTGGCCCAGAAAGAGGTGCCTACGCGACGGGGACGTCCGCCAGGCCTTGAA
GAGCACAAAGATCTCGAACGCTGCCATCATGAGCATGGTGAGGGTGGCCAGGACTAGCACCGGCACCACCCACGGCTCCTTGCGTAGTCCCGCGAAGGGCAAGTGGGTGGAC
ACTATACTGGGCGGTGTGCTGTTGGAGCTGCTCCCGGGAGTCACCCGTCCAGTGGTGCCGGCTGCTCGCGCCTTCGCCCCTGAACTCCGACGGGGCGTGGCTGGTGTGGTGG
TGGACGTGGTGCTGCTAGAACTGCTGGCCGATAGTGTGGGTACCACAAAGAATTCCGTCGACATTTCGAGACGGGCAGCACTGGAGACATTGCCACCCGGAAAATTGAGATC
CAATTGTTTGGCAGATCGCTGGGTCTGAGGAGCTGCCGGAGTAGTGGTACTAGTGGTGCTACTTGTAGTACTTGTAGTCGTGCTGGTAGTAGTGGTGGTGCTGGTGGCATTA
CGGCTACTCGGCTTAAGCACCTGGTAGGCGTTGCGATCCGATTCCTGGCTGGTCACCAGGCGCACCCGCCGCCAAATGTTGTTGCCATTAACGCTGAGGCTGCGCGACGCCG
GCGTGGTGGCGATCAACTGCATGATCAGCAGCAGGATCAGCAGCAGGTGGCGTCTACCGGCCATTGGATGTGGTGACCTTCGATGCCCTGGACCTCTGACAATATAATCAAA
CGCGGCACATCTCTTGCAAACGCACACTAACCACGCACTCGCGATCGCGATTGTTCAGGTTCGTAAATTCTTCAGTCCTCGCAGGCCCTAGATCATGAATTCTTCTTCTCGA
TTCTCGCGCTAAAACGTCCGCTCGCTCCGATCGACCAACTGCAACTGACTGGCGGCTCTGCAAGTTGGGCTTTGGACCGGCTTAGTCGTTTGAGACCAGTTTGACGGCAGCC
AAGTCGAACTAACACTTGTTTGAGTTCGATGTGCGTTTGCTAACCCGGCTAACACTTACAGGAACGCAACATGATGCGTGGGAGGCAGTTGGGGTTTGGGGCGCTTTATATG
ATTTTTTGCAATGTAGAGTATTGCACGAGCATTTACTAAAGCATCCTTTGAATTTGATACAACATTCGTTACTCGATCTTCTGGCTTTAAGGTTTAGTATATAGCTTCAACA
ACTAATGTTCTAAATAGCTCAATGTCTTGCTGAAGGGCTATTGATTTCCTCTCGAAAGCTATGGTACTTAATTTCCTATTTAGAGCCAGTAAGAACTCAATTTATCTTTAGC
TTCCGTGCTAATCCGCTGGAACATATACTTTAACACAACATTTACATACGTAAGTTTTTTTCCATTTGATAATTATTCCAATTAGATTTCATATTGCGGATATTAACATT
GGACACAATTTTTTCCTCTGCGGTTGGTTGCAACATCAACTGCAGACGCGATCCACAGACCTAGCTCAGGATCCACACTCAAACCGACTGCGGGATGGAGAACACTTTCATT
CGTTCTGCATATGCGGAATTTCGGTTTATGGGCTCTGATGCTCGAGTTGGCTGGTAGGAAAGTGCTATGGGCCATTATCTATTAGCCAACCGCTTTTACGGTCTCCTCGCGG
AGCCTCTCTCCTCGAAAACGAGCCAACAAAAGGCCAAATAGCGAAATCCGAGCACAAAAGGTGATTGTAAATCAAAAGAGCGCTAATGGTATATCGATATACAAAGTGTTGA
GGCGGGAAAGGCATCGCCAATAGCGTTAGAAACGGCAATAGCAATAGCAACCCGCTCTGATTCCTTAGGTCTTATTGTATTCGGCGCGCGCTTTCAGTCAACCGCAGTTAGA
ATTCAGATCACCTGAATGAGTGGGCA
(SEQ ID NO: 76)

Exon: 2946..1001
Start ATG: 2752 (Reverse strand: CAT)

Transcript No. : CT22117
TCGATCGGAGCGAGCGGACGTTTAGCGCGAGAATCGAGAAGAAGAATTCATGATCTAGGGCCTGCGAGGACTGAAGAATTTACGAACCTGAACAATCGCGATCGCGAGTGC
GTGGTTAGTGTGCGTTTGCAAGAGATGTGCCGCGTTTGATTATATTGTCAGAGGTCCAGGGCATCGAAGGTCACCACATCCAATGGCCGGTAGACGCCACCTGCTGCTGATC
CTGCTGCTGATCATGCAGTTGATCGCCACCACGCCGGCGTCGCGCAGCCTCAGCGTTAATGGCAACAACATTTGGCGGCGGGTGCGCCTGGTGACCAGCCAGGAATCGGATC
GCAACGCCTACCAGGTGCTTAAGCCGAGTAGCCGTAATGCCACCAGCACCACCACTACTACCAGCACGACTACAAGTACTACAAGTAGCACCACTAGTACCACTACTCCGGC
AGCTCCTCAGACCCAGCGATCTGCCAAACAATTGGATCTCAATTTTCCGGGTGGCAATGTCTCCAGTGCTGCCCGTCTCGAAATGTCGACGGAATTCTTTGTGGTACCCACA
CTATCGGCCAGCAGTTCTAGCAGCACCACGTCCACCACCACACCAGCCACGCCCCGTCGGAGTTCAGGGGCGAAGGCGCGAGCAGCCGGCACCACTGGACGGGTGACTCCCG
GGAGCAGCTCCAACAGCACACCGCCCAGTATAGTGTCCACCCACTTGCCCTTCGCGGGACTACGCAAGGAGCCGTGGGTGGTGCCGGTGCTAGTCCTGGCCACCCTCACCAT
GCTCATGATGGCAGCGTTCGAGATCTTTGTGCTCTTCAAGGCCTGGCGGACGTCCCCGTCGCGTAGGCACCTCTTTCTGGGCCAGATGCTGCTTCTCGGTCTGTTCGCCTGC
GCTAGTTTGGGGGCCATCATCACCGCACAGCCCTCGTTGATCAGTTGCGGGGCCATCCGCTTCGGAGTGGGCGTGGCCTACGCCCTGGTCTTTGCCGCCCTGCTGGTCAAAT
GCGTATTTCTGATTAGTCTAAATGGAGGGGTGTATCTGCCGGCTCCATACCAGGGATTGCTGCTCCTCTTTGCGCTGCTCATCCAGGTAGCGATCGGAGGGCAGTGGCTGCT
CACCCAACCACCGGAGGTATACACCACCAGTGTGCCTGTGATGGGCAGTGGCTTCCTTAGCACCACGGTGGCCTCACAGACCAACTACTCGGCACTGTTTTACCCCACGTCC
TACACAACGCTGGACGGGACTCCGGAGATCTACACCCGGATAGCGGCGGTTAGCACCGTACTAATACCGCTCTGCAAGACGCAGTTCTCGGAGCTGCTCTTCTCGCTGATCT
ACATTGTCTTTCTGATCGTCTTCATCGCCGTGCTAGCCATCAAGTCGCGCGGCATCAGGGACAACTACCGTGAGGCCACCTACATTGGGCTGGCCATCGGAGGAGCCATCCC
CATCTGGCTGGGATGGATGCTATGCGGCCTGGCCGTGGCCGAGCGGCACAAGGACGCTTGCGTGGCCTTCGGTCTGGTAGCCACATCCGCCACGGTCTTCTTGGTGATGTTC
ATGCCCAAGGGCAGGCAACTGGCGGCCATGGGCAAGGAGGGCCTGTACGTGGAGGACCGCGAGGAGCAGTTCAGTTCCCTGAGTCGCGCCGGTTCCGGCTACTCGCCCTCGT
TCTTCCACTTCAAGCCCATCAAGTACGGCGTGATGAGCGGCTGCGGACTGCCCAACTCTGCATCCAACACGGGCCAGGGACTCAGCTCCAAACACTGCTCCAGTGCCAACAA
TGGAGGTGGTAAGTCCCGCTCATACGAGCTTTTCCCCTCCAAATTTGTCACATTTTCTGCCCGAGGCGGTCGGACCCAACTCAAGCTAGATACAGTTCCATTTCGGCCCCAT
TCCTCTGCCCTTGTCAGCTGTCAGCCTCTGTTAATTTTGTAA
(SEQ ID NO: 77)

Start ATG: 195 (Reverse strand: CAT)

MAGRRHLLLILLLIMQLIATTPASRSLSVNGNNIWRRVRLVTSQESDRNAYQVLKPSSRNATSTTTTTSTTTSTTSSTTSTTTPAAPQTQRSAKQLDLNFPGGNVSSAARLE
MSTEFFVVPTLSASSSSSTTSTTTPATPRRSSGAKARAAGTTGRVTPGSSSNSTPPSIVSTHLPFAGLRKEPWVVPVLVLATLTMLMMAAFEIFVLFKAWRTSPSRRHLFLG
QMLLLGLFACASLGAIITAQPSLISCGAIRFGVGVAYALVFAALLVKCVFLISLNGGVYLPAPYQGLLLLFALLIQVAIGGQWLLTQPPEVYTTSVPVMGSGFLSTTVASQT
NYSALFYPTSYTTLDGTPEIYTRIAAVSTVLIPLCKTQFSELLFSLIYIVFLIVFIAVLAIKSRGIRDNYREATYIGLAIGGAIPIWLGWMLCGLAVAERHKDACVAFGLVA
TSATVFLVMFMPKGRQLAAMGKEGLYVEDREEQFSSLSRAGSGYSPSFFHFKPIKYGVMSGCGLPNSASNTGQGLSSKHCSSANNGGGKSRSYELFPSKFVTFSARGGRTQL
KLDTVPFRPHSSALVSCQPLLIL*
(SEQ ID NO: 78)

Name: Metabotropic glutamate receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384617
AAATTTGTATAACTTTCAGTTCTTATACTCAGTTATTCCATTAATCTCTAATCCCACGACATTTAAATACACATACAAATTTTTTATACAAAGTTTGCTTTATTGACTATTT
TTGAAGCACATTTACCATATCTGTATTATTAAAATACTTAGCATAGATTACAGAATTTTCAGCTTATTCCATGAAATACTGTTTGTCAAACAAACAGATTTCATCAGTTTTT
GTTTTTATAATCCTTTTTTTTTGGGAAAGATTCTCATATATTTACAATTATAATTTAAAATAGTCTCTACTTTAGCTTACTACATATTCAATTAAATCTCAGGAGGCTGACT
TTTCCGCGGACTTTTATGGTGGCTCTGACCCAACTGGATGCATCCTGCGCCTTTCCTCGCAGAAAAGTATTTTATTTCACATTTTACGGGCGTCTTGGCTGGCTGTCTTACC
TCTGGCTAAGATGCTTATTTTATACAAATCCGACTTTTGCGATTTTAATTGCATAAATCAGCTTAGTTATTTAATTGGTTTATTAATCTACGGGAGAAGCAACCATTTTATT
TTAGGTAAGCACTAAGAAAAAGATTGGATCTTAAGATTCCGATATTATTATATATCTTTATTTACTTAGTAAAAGGATGGGGATTTACGTATATTTTATTTAAGAGGGTATA
TTGGCAGTAAAATATTTATTCAAATTGTGGGTTGTACAGTTTATATTAATTATTAATATGTTATTAGCTGTCAAGGCAATAAGAAAATGTAATCCATTTCAGCCCCCTTCAG
TTAAATATATTGCTAACATATCGATGTGCGGTTGTACAGACTCATTTTCAAATCTTTGAATTTTAATTCTGATCTCGGTTTTTATTACGTAGACACATATTAATACTGCTTA
GTCTTAAGTTCTTCATTCATATTTGGATAAGAAAGGATTATAAAATCATTTAAATGCATACTATAATATTGGCTTATTTAAAGGAAATTCCTACTTAAACTGGGTTACAAAT
CTGTCTGCTGCACTGCGCCCGATCCTCTTCTCGCGGTGGTGCAAAATGCCTCCTGCTGCTCGCTCAAACATATAATCAAATTCTCCGGCGGAGCTGTCACTTTGCCATTGGA
ATTGCATGATTTCTGTGTCCCCGTGGTTCCGGTCTTTTCCGCTGTCGTGGTCGAGGATGTGGTGGTCGATGAATTGTTGTTACCCGCATTCAGTGGCAGCAGTGGAGGAATC
TGCGGATTGCTGGTCAACAGGCGCTTGGAGCCACCCCTCTTTTTACTGCCCTGTTTGGTGAAAACACTGGGCTCCAGCTGGAGTTGAGCGTTGATATCCTGCTTATTCATAC
AGGTAAAGGCCTTGAAGAAGCTCTTCCGGAAGTTCTCACTTAGGAAGGCATAAAGTATGGGATTCACCGCCGAATTCGAGTAAACCAGTGCCCCCAGAAGTAGGAAAATGAG
TATTTCCAGTCGGGAGAGGTCCCTTTGCGCGGGATTCGAGTGAATCAGGGCCACCTTGGAGATCCAGTGGTGTGTAAATAAAAAATCGAATATACTTTATAATTAATTACCT
GAGAAATCCAGTGAGGGAGCCAACATAGAATGTATACACTTATCACCTGCGGATTGAAAAGAAGGTGGGAATCTACAAATGAATTGAACTGATTATTTCGGCATTATTACCG
TCAGTACCAGTCGAGTGACCTTCCTGTGAGCCCGCCTCTTCTCCTTGGACTTCGTTCCTGGTTTGGGACCCACCGATCGCAGTTTCCTTATAACCAAGTAGTAGAAACTCAG
GATAAAGCACAGCGGTGTGGCGAATCCTAGGAAAAATGTGTACAGTATGAAGGTGGTGCCCGAATGCTTCTTGTACGCATCTGGCCACATTATGTTGCACGAGTAATTGATG
CCATCCTCCTGCTCCACAGTGCTGGCATAAAGGATCACGGGCAGCATGAGGACCGCTGAAGTTGACCAGGCAATCGCTGAGACCACTTTGGCAATATGCAGAGTTCGATATC
GTGGCGAGGAAATCGGGTGGCATACCGCTATATATCGATCCGCGGACATGATGAGCAGAAAAATCGACGAGGTGAAGGAGGTGATGGATGTGCTCACCATGTAGGCTTTGCA
CATAAACTCCCCGAATCGCCAGCTGCAAATTCGCATTGTGTACAGCAGAAAGGGTATTCCAATCAGGAAGCACTCGTCTGCCACCGCCAGATTCAGGATATATATATTCGTG
ACCGTTTGCATTTTGGAAAAGCGCAACACCACGTAGATCACCAGGGTGTTGCCAAATAATCCGATAATGCACACAAATCCGTAGAGCACCACAGTGAACAAGTCAGCAAAGG
AATTCCGCGTGGCTATGCAATGTTGATAGGTGGGCAAATCAGTGCCATAAAGGGACTCCTCTGGCTGAATTGTGGAACTACCACTAATCCATCTATGGTTCAGTTCCGTGGT
ATATAAGCTCTCGTTCGTGTTATACCAATTTGTGGCATTTGCCTCTGAATCGGCTGTCATTTCCCCTTTCACAAACTGGAGGACATCCATCATCAGCCACGTAAACATATTT
CCACCTAAGCGCTGTCCAGTTCAAATTTAATTATTTTAAATAGGTTTTACGCGTTGTTGCTCCCCGCTCGTTTTGTTCTTGCATTTTGGCGTTGCTGCAAAATACATTTTGT
TATTACGTGAAATATTTTCAGACACAGTTGTTTTGGCCAAACTTGATTAGTTACTAATTGAAATACTGTGGTGCTATTAAAGAAAAAAACCTAAAACCTTCACTAGTCTAGT
CAATATTTTTTATTCCAACGTAAGTTTAAAATGTACTGTAACAATTGGTGATAAGTATTCAAATGGAATTAAGTTAAGTCTACAAGTTTTATAAGATATTTTAATATAACTA
GATTGCAGCTCTTTAATGCATAGTGTATATTTTTAATATATTATGAGTTTACTTTTAAGTGGCAGTCAATTTTATTTGCTGTCTTAACATCTTTGAATCGTATGTACCCCAT
TCGTAATGCCACTTCACATACAGATTCTTTTGCAACTTTACTACAACTTTTACTGACTTTCTGAGAAATCCCCTCGAAAATTGAAAACTTTCCTGAACAACTTGTAGAAAAA
CAAACCTTTAAAAATACGACGTATGAAATAAATCGTCTGTTTCCTGCTTTTGAAAAGTGAGTTACCACAAATTTGGTTTGTAAGGCGAAGGAAATTAACATTTATATTTAAA
TTAAATTGCTCACAGCTTGACACGTCCAACAAACAAGAAACAAGAACTTTTTTATATTCTGACCTAACTCGACACTTTCTGTAAATAAATTAGGCGTAAGGAAAAGCTGGGC
AGCATTTATGCAAATATCAGGTGAAAATATTTGTCTTCTCGGAAGTCGGAAAACTCCAGCTATACAGATCCCCATGGTTTTGTATTAGGAGATTGAAATTTAAGCCGATTTG
TTTGTGCATTCCATGCACATCAAATATTGTTGACTGAGCCAAAACACATGTG
(SEQ ID NO: 79)

Exon: 2524..1679
Exon: 1510..1001
Start ATG: 2524 (Reverse strand: CAT)

Transcript No. : CT22465

ATGACAGCCGATTCAGAGGCAAATGCCACAAATTGGTATAACACGAACGAGAGCTTATATACCACGGAACTGAACCATAGATGGATTAGTGGTAGTTCCACAATTCAGCCAG
AGGAGTCCCTTTATGGCACTGATTTGCCCACCTATCAACATTGCATAGCCACGCGGAATTCCTTTGCTGACTTGTTCACTGTGGTGCTCTACGGATTTGTGTGCATTATCGG
ATTATTTGGCAACACCCTGGTGATCTACGTGGTGTTGCGCTTTTCCAAAATGCAAACGGTCACGAATATATATATCCTGAATCTGGCGGTGGCAGACGAGTGCTTCCTGATT
GGAATACCCTTTCTGCTGTACACAATGCGAATTTGCAGCTGGCGATTCGGGGAGTTTATGTGCAAAGCCTACATGGTGAGCACATCCATCACCTCCTTCACCTCGTCGATTT
TTCTGCTCATCATGTCCGCGGATCGATATATAGCGGTATGCCACCCGATTTCCTCGCCACGATATCGAACTCTGCATATTGCCAAAGTGGTCTCAGCGATTGCCTGGTCAAC
TTCAGCGGTCCTCATGCTGCCCGTGATCCTTTATGCCAGCACTGTGGAGCAGGAGGATGGCATCAATTACTCGTGCAACATAATGTGGCCAGATGCGTACAAGAAGCATTCG
GGCACCACCTTCATACTGTACACATTTTTCCTAGGATTCGCCACACCGCTGTGCTTTATCCTGAGTTTCTACTACTTGGTTATAAGGAAACTGCGATCGGTGGGTCCCAAAC
CAGGAACGAAGTCCAAGGAGAAGAGGCGGGCTCACAGGAAGGTCACTCGACTGGTACTGACGGTGGCCCTGATTCACTCGAATCCCGCGCAAAGGGACCTCTCCCGACTGGA
AATACTCATTTTCCTACTTCTGGGGGCACTGGTTTACTCGAATTCGGCGGTGAATCCCATACTTTATGCCTTCCTAAGTGAGAACTTCCGGAAGAGCTTCTTCAAGGCCTTT
ACCTGTATGAATAAGCAGGATATCAACGCTCAACTCCAGCTGGAGCCCAGTGTTTTCACCAAACAGGGCAGTAAAAAGAGGGGTGGCTCCAAGCGCCTGTTGACCAGCAATC
CGCAGATTCCTCCACTGCTGCCACTGAATGCGGGTAACAACAATTCATCGACCACCACATCCTCGACCACGACAGCGGAAAAGACCGGAACCACGGGGACACAGAAATCATG
CAATTCCAATGGCAAAGTGACAGCTCCGCCGGAGAATTTGATTATATGTTTGAGCGAGCAGCAGGAGGCATTTTGCACCACCGCGAGAAGAGGATCGGGCGCAGTGCAGCAG
ACAGATTTGTAA
(SEQ ID NO: 80)

Start ATG: 1 (Reverse strand: CAT)

MTADSEANATNWYNTNESLYTTELNHRWISGSSTIQPEESLYGTDLPTYQHCIATRNSFADLFTVVLYGFVCIIGLFGNTLVIYVVLRFSKMQTVTNIYILNLAVADECFLI
GIPFLLYTMRICSWRFGEFMCKAYMVSTSITSFTSSIFLLIMSADRYIAVCHPISSPRYRTLHIAKVVSAIAWSTSAVLMLPVILYASTVEQEDGINYSCNIMWPDAYKKHS
GTTFILYTFFLGFATPLCFILSFYYLVIRKLRSVGPKPGTKSKEKRRAHRKVTRLVLTVALIHSNPAQRDLSRLEILIFLLLGALVYSNSAVNPILYAFLSENFRKSFFKAF
TCMNKQDINAQLQLEPSVFTKQGSKKRGGSKRLLTSNPQIPPLLPLNAGNNNSSTTTSSTTTAEKTGTTGTQKSCNSNGKVTAPPENLIICLSEQQEAFCTTARRGSGAVQQ
TDL*
(SEQ ID NO: 81)

Name: Somatostatin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384531
AGCATGCGGCTGAACAAAAAAAAACTCGAGGAATGGCCGCAATTTCGGTGAGCCATATAAACAACATAACATATAGTTTTCCCATCCATCTACGCTGTCTGCGCCACTGGCC
AAGTCCAAGTTGTCTTGACGGCCCGGCAAAAAGCAACATCCAAGGAAAATGTTGCCATTTTAGATGCGCCGCCCGCTGAGCGGAAAAGTCGCGCGGAAATTACTAAATCGAT
ACGTGCTGCGTATCAGATTTTCCATGGGTTCTGCATGCAAAAATACCTTGAAATCCTTGCCTGCAGTCGCACACATACACGTATTGCCATCTTTGTACAGACAAGCTTGACT
AAGTCGAACTTTACTATGTTATTCTACTAAAAGGTTTTAAAGCAAATGATCTATACTATACTATACTATTAACTTCAATAGAATTCCTATTTTCTCTGGCTAGCAAACAAAT
TTGGATCATAAAAAGATTTTCATGGCATTTATGAGATACGTAGATTGTGTGGCTCAACCTTTTATCTATATTATTTCCATAAATTTAAGCACCAATGGAAATGTTTATCTCC
AAAGCTTAGGAAAATTTGGAAGTATATATAATAAAAACCTTTTATGTCTGCAAATAAGAGATATTACGTGTATACTGAGTTTCGCTTGAGTTTCGGAGTTAAGAGGCTTCCC
TATAGACTGTTTGAAAATCCTATAATGCCTTGGCATTTCTAAAAGGTCAACTGAAAGCAGTGGAATTGCTGGTGGGCGAGGTTCGTCTCCTAAGTGGCAAAATGATCCAGAA
TCGGAAGCAAACCCTGACGATGAAATAAACTATACGAGAAACTGGAGAACAATTACACATGGACTGTGGGTGCAATTAGGGACCACACCTTTTTCGCTGCCGACGCGTCTAA
AGTTAATGAATATGTTAATAATATGTTATGCGCGTGCAAATTTGTTCAGGGCTTGAGGGGTTTTCTGTGTTTCTGTTTCCAGGCGAGTCCAGTCCAATCCAACTCTATCGCC
TGATGGACATGGTGTTCGTGGGGAACTGTGGCTTCTGGCCGGCGCTGAAGAAGCGTTTGCAGGTGAGGCGCCAGAAGGCCGCCCGGAAGTCGACGCTGTAGAAGGCGTAGAT
GAAGGGATTAATGGCGCTATTGAACCAACCCAGCCAGGTGAGGGCCTTGGCCAGCATTTGGCTGGCCTGGTGTTCGGCCAGGAACGGGGTGATCAGATAGTTGATAAAGAAG
GGCAGCCAGCAGGCGATAAAGCCACCCACCACAATGCTCAGGGTTTGAGTGGTCTTGTTCTCCTTCTTCAGCGACGTAATCCGATTGGACAGCGATTTACTGTTGGTGTTGG
CACTGGTGTTCCCGCTAGTCGATGTCGTGGTTACGCCCGCCTGGTGATGATAATGAGGATGGTGGTGATGGCTTCGCAGTCTCTCGCGGTCCCGCTCCCCATGCGAATGCCT
AAAGCTGTGTGTTCGCGGTGGCTGCGATAGGGAGTGGGTCTGCACCTGGCCCGTTTGGCCCGGCGGCTGGGACTGCCATTGCGCATCCAGGAGACTACTGCCTGGGCCCATG
CCACTCGTCATTGTGGTAATGGTCGTTGAGGCGGTGGAAGCGCGCTTCAGCGAGGAGGGTCGTGTGAGTTCGTAGCACCCATTCTTGGCGCAGCCCATGGAGCCACCTACTC
CGCCGGAGGGAAGGGCCAGCAGGGATTGACAGTGTGTGCCGCCAGTGGCACTACTTGCTCCTCCACCCCCACCACCTGCTCCTCCAGCGCCCATCGCCGCACAATTGTCGCT
GTCCGCTGAGCATCATGTCTTGGAGCTCCTTGGCTATCGTCTGGTTGGAGAACGTTCGCGATGTGGCCTGTCTCTGGCGCTGGCCCACCGAGCTGTGCTGCTCCTGCTCCGAC
AGCTCATTCTCCACGTCCGCTGCCGTGTAGCGCTTGAATTTCTAGACACGCAAAGTGCACAAGGCGGTTAAGTCAATGTCGATGATAAATAACTAAAATAGAGTCAATGGGC
GGCAAAAGGCGGCAAAGATAAAAGAAACGGAAACGGAAGTCACGCTAGGAAAAGCCGAAGCTTTAAACACTCTTGTAGGCATCATCATATACTTTAAAGGGAACCACACTGG
TAAGCAGATGGCTTTAAGCTGATTTAAGACTACATACCCAAATGCCATGCGGCCATTGTTCTATCAACTGCACGATGTGCTAATCTCTCTAAACATAAAACTGAATTTCAGA
CCGGTTTTAATGCGTATGGAGCTGCCTGGACTTGATTCAGTTTAAGTGAATTTTAAGCCTATTTTAAAATGCACACCAGGTACACATACAAATGTATATATGCAGATGCAAT
TTTCGAAGCGAGCATCGTTTAGTGTTAGCTATTTAGTGATGCTAACTTAAAGCTCGACATAAATGAGTCGAATATGTAATTACCCTGGTATTCAAGGATGTAGCAACATTAG
CAACATTTCATAACTAACATTTTTGAAAAGTATTCAACTCTAAAGCGCCAAACTGCTCTCTCGTTAGAAAACGGTTAGTTGAAAACTCATTAATTCAGCTATACGGGCTGCC
AGAGTTGAAGTGAAAAACTTGTTTATAGCGGACACTGGACCGGAAATTGGCACATACCCCGATGCGAGGGTATGGTATTTAAATCTGTGTGAGCTCCTGGCATTTGGGAAT
TATGAGGCAAACGTGCCACAGTTTGTGCCGCACGTGGAGTGTCCTTTAAGCCGGAGTGGTTCGCAATAGAGATGACTTTCGCTGGTCGGGCCAGCGTTTGTTGTTTGCAGCT
GCCGTGGTTTTCCTATTCCATATTCCGTATTTGTGCACTATGTTGGCTTTGGGGTTTTGGATGGCGGGAGACGTGCGAATGTTTGCAGGCAAACAGAAGGACATGCGGACAT
TGGGACATGGTTGCGGATGGCTGCGACGGAGCTTACCTTGTTGTGAACACTTATGTCGGTCATATTGTCGTGCCTGGATGCAATGACACAGGAAATTCTTGCATACACATAA
ATCATGACTGCCATGGGTATGAAGAAGGAGCCCATGGCCGAGAAGATGACGTAGCCCTCGTTCTGGTTGTATCGGCACTCGCGCAGATCCCTTCGTCCTGGCTCGTACCTGC
GGGAAATGGAAGTGAGCAGGATACGCTTAATGGGCCACCAAGTTAGTCAACTAATGCAGGAAGTTCCCAATTATTTGACTACGGAGCTGTGACCCTCGGAGGGGTCAGAGGT
CGCAAGAGCAGCAGACAATTGCCAGCTAATTATATTCGTTTGTTCAGCTGGGTGTCGCCAGTCAGTCTGCTCGGAGGTCTCGAAATTTATTCTAATTAATTTACACTCACAC
TCGGCGACATTTGTAATGCTTAAAATGGACCAAAAAGATACCCAAAGTGGTCCTTCAAAATTATGCAATGGCATCTTTTGACACAATTTTCTGCAGTCAACAAAATTGTTGA
TTTAATTTCATTTAAGCAAATTAAGCTTTAATTATTTTTGTAAGTTAAATAAATTAAACATCTCTTGGTTACATTTCGGAATTCTGCTGCCTACAGGAACATTTAGTATATT
TCTTCAAAAAATCTGTTGACTGTGCTAACGAAGTTAGATATAAACAAAGCACTTTAATTAATTCTCCTTTTAGTTAGTTAGACAAAAGTAGAGAGAACAAAATTTGTTCTTA
AAGCTAGCAGATGCGGTATACCAACAAGGAAGTTGATGTTTAGGTTAAAATTCGGTGCATCTATTTAGTGTGTGGACTGTTGACCAACTTAGAAGACTTCTTGCTGCACTCC
TTCTTCGTTATGGAAATAACCCTACATACTGATTATTTCTTCAAGTGTACTTACCATCCCAGCATGGGGGGACAGGTGATGGCCAACGCGAGCAGCCAGACGATTAGGATCA
TGATGAGGGCCAATCGTTTGGAGCGCCGTTTACGGGAGTACGTGAGCGGTCTGGTCACGGCCAAATATCTGTCGGATGGACGTGAAGGATGCAAAGATTTACGATATCAGGG
GAATACTATGGATGTTTTGGATAGGCCTATGCTCACCTGTCCACACTGATGGCGCACAGGCTGAGAATGGATGCCGTGCAGAGGAGCACGTCCAGGGAGATCCAAATGTCGC
AGAGCACCCAGCCCAGTTGCCACGAGCCTGCAATCAAGATTCCAAATTAAAAACTTTCCGCCAGCCTTTTGCAGCGCGGAGTGGGGCAGCCAACTTAAAATTCCGGTCCAAA
GTCAATTTGATTAGCAACATTTTCCATGGCTTTCAGCAAATGCCTCACCTAAGGCCAATTAAAGCGGCCAGCTCGGCAAATTTTGGCAATGTTTCAGGTCCCCTTCAACCCG
TTAGTGGTTTAATTCTATACCGAAAAAGCACCGATTTCTCCATTTGACCCCAACACTTCGAGCTTCTACAGTGCTTTCTCCTCGTCGGTGGCTTGTCAAGCGATTAATTTTT
AAATCAAGTGCCCCTTAATGGTCTCCCGGCGCCAAATGAAAAGTTGTTTTTGTGAAAGGCGCTAGGAGCGAAGGAACCAAGGAGCTCATTACACGGATATTGCCTATGGGCC
AGCACACTTGGGTTGGTCTAGTTCAATGTAAGGTTGCTGAATTTGCTGAATCTCCAAGGACTCGGATTGAAGCAGTTAAAGGCAAAATTTGGAGTGAATTAGAAAAACACAA
GGATTACTTAATTAATTATGGGAATGCTTGTTCAGTTGTAGTGTATGTGTATGTTTTTATCTAAAAACTGGAGTAGGAACTGGCACAAAGCAGTCGTCATATCAGCACTTTC
AGTGCCTACCCCAATTTCAACTGGATGATAAGGCTCCAAAAAGCAAACTATAAATCCCGACCGTGTGGGTGTATTACGTCCTAACTTAATGATATATGTGCGTATACTTTGG
ACAAAAATGAACTGAATGTGGGGCTTGGTAAATTGTTTGCCTCGAAAAAACCTGGGGGCATGAGTAATTATGAATGAGATTCAAAAGGTATCTCCACAGAGAGAAACTAGTG
GATGAAAATCGGAAAACAAGCTTAATGTTCCATAAGACAGAAGCCAAGTGAGCTCAAATCAGAAGGGTTTAGCATATTTTCCATTCAACTTCACGAGTTCAGCAACAAAGAT
CAACTATTTTCTGACGAGAATAATTCATTCCGATTCGGATTGTCAGTGGTCCGGATGATTATGACAAACTTTCGACATGCCGTTGCCTTTAAAGTATTTCCACAAAGTTGGG

CTCGGGAATTTCCCCTTTTTCGGCTTTTTTTTTCGTGCTGAGGAAATTGCACGAGAGCCGAGGCAAAAAAATGAAAGAAATTCTTTCAGTTCAAAGAATACAAATCCATAA
GGCGAGAGAGATACATACTCGCATATTTCACCTAGCAGCATTTACATAAATACACGCCCCCTTTGCCCCGCCCCACCATCTGACACACCTCCTCCTCGAGTTCCTGTTCGTG
TCCGTGCTGCTATTCCTGTTCGTGCTACTTTGTTGGGATACACTTGGAGAAAAATATATATATATGCTAGTCCTCCTAGAAATCCCTACACATTGCTCTAAACAGAATCTCA
TTGATATTTACAGGTTCGAAAGTTCCTTAAATACCTCAAGAAACATTGCTACTACATTTTTACTTCATGAGATACAAATTACAAAGACTTGTATTGGATTTGTTCTCAGTGC
ATTCGCTTGGTTCTGGGCGCCAAGCTTTTGTTAGCAATTTTTTGAGGGATTTACCGGAATGGCATCTTCATCTTGTAATCACGCGAAAAGCCCGGCTAGAGGCGCCTCCTTT
CGCGGAAGGACTCTAATGCCTTAATGGCGCTGCCACACACTCACACACACACTTACACACTCACACTCATACACACATAGACTCACCTATGAGATGGACGGCGACGGCGGGG
GGCATGACGAAGATGCCCACGAGAAGGTCGGCCACCGCCAGGCTCATCACAAAGCAGTTGGTAATGGTGCGAAGCCGGCGAGTGGTCAATATGGCCAGTATGACCAGGGTGT
TCCCGATGACGGTGACCACGATGAAGGTGGCAAAGACTGCAATCAGGATGATGCCCTGCCAGGAGAGCACCAGATCGTAGAAGTGCGCCCATTCGGCGGCCGCCGCCGACTC
GTTGTAGTAATTAAATGTGGTCACGGGCGCCGTGGGTCCGGGTTCGGTGAGAGCCACTGCATCCGTACCGTTCGGTCCGCTGGGATATCCACCTCCTACTCCTCCTACTCCT
CCTCCACCTGTCCCATTGCCCGCATCCTCATATACGATGGATATGTTTTCCGGCAGACCACCCACTCCCACTCCCGATCCCAGTCCCAATCCCCGCAGGGAGGAGGAGGAGG
CAGCCACATACATGAGCGACGGATAAATGGCGCTTAAATTATTCAGCTGCAAGTTGACAGCCATGGATCGGCCAGTTGGGTTCACAATCGAGTGTGGCTAGCTCCGGCAGCA
GGCCGCCTAGGCCCATCCATATCGTGCCATATCCCTGGAGTTGACGGAGCAGGAACACGGCCCAACTGTGGTCATTGTGTCCAGTTGCACGTCGGGTTAACTGGAATTAAAG
GAAAGTTCGACATTTAGCAAAGTCCTTTGCATGGTAATAAAAACGAAGTCAAGAGGAAACAATAACCATTTGCATATCCCTTTTATGCTGAAATCTATTTGTGTTTGTTTTA
ATTTTTTATGTATGCACGCATTATCTGATTTATTCCTTGTGACAATTTCCGAAATTAAATTGATTTCACAGCATTCCATACAATCTCATTTAAATTGCAATCGTCGAAATTG
TCAATTTCATTCATTCGCTTGCAATTAAGCATTCAACAATCTCAAGGCAAAAAGTGCACAGCATTGAATATACCCGTCTGTGTAAATTACGTTGAAAGTTGTCAACGGTGTC
AAAAGATTCTTAAGATGGCCTTTGTATTGGGTTGCCATTACCGCGATACAGTTAAATACTTTACTTTCTAAGTAGATAAACGTTAATCCTTTAGTTAGTCATGCTTTCAACG
AAAATGCCGAAGTGCGCTACTGTGAGCAATAAATCAAATTAAGTTTACCTTATCATCTAGAAGCACCCCACTGGCTTGATTTAATTTAAAATGTATGATTCCATTTGCTTTGG
ATTAGGTACCGAACTTAGAGACTTAAGAATAAATAATTCGTGGAGGTATTTTCTATCAGAAAATTAAAAATTTTTATACAAATATTCACGGTTTGATTTGGTGAAGAAAATT
CAAGAAAATCATCAACTTCGGAATTCCTGGCGCATTCTAATAATCATTATATGAAAATTACAAATATAATGAGCTTTATACAAAGTAAGCACTATTACTACAAAGTGGCGAA
ATGTATTACCAATAAATTTGCCAAGGGAATGGCTACATTTTGGGAATATTTTATTT
(SEQ ID NO: 82)

Exon: 6448..5911
Exon: 4171..4069
Exon: 3988..3863
Exon: 3132..2949
Exon: 1945..1001
Start ATG: 6448 (Reverse strand: CAT)

Transcript No. : CT22855
ATGGCTGTCAACTTGCAGCTGAATAATTTAAGCGCCATTTATCCGTCGCTCATGTATGTGGCTGCCTCCTCCTCCTCCCTGCGGGGATTGGGACTGGGATCGGGAGTGGGAG
TGGGTGGTCTGCCGGAAAACATATCCATCGTATATGAGGATGCGGGCAATGGGACAGGTGGAGGAGGAGTAGGAGGAGTAGGAGGTGGATATCCCAGCGGACCGAACGGTAC
GGATGCAGTGGCTCTCACCGAACCCGGACCCACGGCGCCCGTGACCACATTTAATTACTACAACGAGTCGGCGGCGGCCGCCGAATGGGCGCACTTCTACGATCTGGTGCTC
TCCTGGCAGGGCATCATCCTGATTGCAGTCTTTGCCACCTTCATCGTGGTCACCGTCATCGGGAACACCCTGGTCATACTGGCCATATTGACCACTCGCCGGCTTCGCACCA
TTACCAACTGCTTTGTGATGAGCCTGGCGGTGGCCGACCTTCTCGTGGGCATCTTCGTCATGCCCCCCGCCGTCGCCGTCCATCTCATAGGCTCGTGGCAACTGGGCTGGGT
GCTCTGCGACATTTGGATCTCCCTGGACGTGCTCCTCTGCACGGCATCCATTCTCAGCCTGTGCGCCATCAGTGTGGACAGATATTTGGCCGTGACCAGACCGCTCACGTAC
TCCCGTAAACGGCGCTCCAAACGATTGGCCCTCATCATGATCCTAATCGTCTGGCTGCTCGCGTTGGCCATCACCTGTCCCCCCATGCTGGGATGGTACGAGCCAGGACGAA
GGGATCTGCGCGAGTGCCGATACAACCAGAACGAGGGCTACGTCATCTTCTCGGCCATGGGCTCCTTCTTCATACCCATGGCAGTCATGATTTATGTGTATGCAAGAATTTC
CTGTGTCATTGCATCCAGGCACGACAATATGACCGACATAAGTGTTCACAACAAGAAATTCAAGCGCTACACGGCAGCGGACGTGGAGAATGAGCTGTCGGAGCAGGAGCAG
CACAGCTCGGTGGGCCAGCGCCAGAGACAGGCCACATCGCGAACGTTCTCCAACCAGACGATAGCCAAGGAGCTCCAAGACATGATGCTCAGCGACAGCGACAATTGTGCGG
CGATGGGCGCTGGAGGAGCAGGTGGTGGGGGTGGAGGAGCAAGTAGTGCCACTGGCGGCACACACTGTCAATCCCTGCTGGCCCTTCCCTCCGGCGGAGTAGGTGGCTCCAT
GGGCTGCGCCAAGAATGGGTGCTACGAACTCACACGACCCTCCTCGCTGAAGCGCGCTTCCACCGCCTCAACGACCATTACCACAATGACGAGTGGCATGGGCCCAGGCAGT
AGTCTCCTGGATGCGCAATGGCAGTCCCAGCCGCCGGGCCAAACGGGCCAGGTGCAGACCCACTCCCTATCGCAGCCACCGCGAACACACAGCTTTAGGCATTCGCATGGGG
AGCGGGACCGCGAGAGACTGCGAAGCCATCACCACCATCCTCATTATCATCACCAGGCGGGCGTAACCACGACATCGACTAGCGGGAACACCAGTGCCAACACCAACAGTAA
ATCGCTGTCCAATCGGATTACGTCGCTGAAGAAGGAGAACAAGACCACTCAAACCCTGAGCATTGTGGTGGGTGGCTTTATCGCCTGCTGGCTGCCCTTCTTTATCAACTAT
CTGATCACCCCGTTCCTGGCCGAACACCAGGCCAGCCAAATGCTGGCCAAGGCCCTCACCTGGCTGGGTTGGTTCAATAGCGCCATTAATCCCTTCATCTACGCCTTCTACA
GCGTCGACTTCCGGGCGGCCTTCTGGCGCCTCACCTGCAAACGCTTCTTCAGCGCCGGCCAGAAGCCACAGTTCCCCACGAACACCATGTCCATCAGGCGATAG
(SEQ ID NO: 83)

Start ATG: 1 (Reverse strand: CAT)

MAVNLQLNNLSAIYPSLMYVAASSSSLRGLGLGSGVGVGGLPENISIVYEDAGNGTGGGGVGGVGGGYPSGPNGTDAVALTEPGPTAPVTTFNYYNESAAAAEWAHFYDLVL
SWQGIILIAVFATFIVVTVIGNTLVILAILTTRRLRTITNCFVMSLAVADLLVGIFVMPPAVAVHLIGSWQLGWVLCDIWISLDVLLCTASILSLCAISVDRYLAVTRPLTY
SRKRRSKRLALIMILIVWLLALAITCPPMLGWYEPGRRDLRECRYNQNEGYVIFSAMGSFFIPMAVMIYVYARISCVIASRHDNMTDISVHNKKFKRYTAADVENELSEQEQ
HSSVGQRQRQATSRTFSNQTIAKELQDMMLSDSDNCAAMGAGGAGGGGGGASSATGGTHCQSLLALPSGGVGGSMGCAKNGCYELTRPSSLKRASTASTTITTMTSGMGPGS
SLLDAQWQSQPPGQTGQVQTHSLSQPPRTHSFRHSHGERDRERLRSHHHHPHYHHQAGVTTTSTSGNTSANTNSKSLSNRITSLKKENKTTQTLSIVVGGFIACWLPFFINY
LITPFLAEHQASQMLAKALTWLGWFNSAINPFIYAFYSVDFRAAFWRLTCKRFFSAGQKPQFPTNTMSIRR*
(SEQ ID NO: 84)

Name: Octopamine receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384651
GTGCTCGACATTTCGCTTGCGTCCCGTAAAGAAATTGTCCACCACGATGACCTCGTGCCCCTGGACCATCAGATCGTCGACCAAGTGGGAGCCCACGAAGCCAGCTCCTCCG
GTAATTAGTATGCGTTTGCGATTCTTATAGTTCAGGTATTTAACCTTCGGATACTTGCGCGGTGTGCTGGTCTGGAGGCTACGAACTTGTTCCTCCAATCTGGCCAAATTCT
CCCTTGTGCGCTGGAGTTCGGCCTTTTGTTCTCGAATAAGAGAGCTCTGCTCCTCGTACGCCATCTGAAGGCTCCTTTGTAGCGGACTTTCCGTGGGCGGCCATTTGGCCTG
AACCTCCTCGACGCCGGGAACGGAAACGGCGACCTTTCCGCTTGGGCAGAAGCTGGCCATTCGGTAAAGGTAGACCAGGAGCAGGAGGAGCAGCGAGATGGCTGCAACGATC
TTCAGACGCTTTTTGGTGGCAGTCATCGTCGGTTATGGGTTGTGTACTGCAGGGTTCCAGTGCTTATCAGCAACTGATATCGCCACATTGCATTCTCGTTCGGCTTTAGAAA
CTATATAAACAATACAATTGCGTTATCAGGAGGTGCCAGAGTGGCCAGATGGCTCAAACGGAATTTCTAAATGGATTAATAATCTAACGCCAGGGAAAAGCTCCCGCCTAAT
TTTAAAATTTACATCAAAAGTATCGACAGGCGAACATTTTGATTTTGGATTATAAATTATAATGAATTGATAATGAATTAGACATTTTTGTTGGTGATCTTCCCCAGGGACG
AATATTTTTCGTTTTAACGATAAAAGGTCGGAAATTGCCAAAAAATGAGTTTGAAAACTAAATTTGTTTGAGGGCTTTCTTAGGCTGGTCAGCAGGGACTACTTTCCGAATA
AAGTGTTTGCCCAAACGGCTATTGAAGGCTATTTAAACTGCCAACGATGAGCATCGATCGCCAGTCGGGCAAAGTTCTTAGTGCCAAGTGTTGTAGTGCTCAGCATGAGGCT
TCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCAAATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACATTGAAAGG
CAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTACGAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAAGCATTTAA
GGCCTGTGTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTTGGCCAACGGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCCCGGTTCAA

ACCCTATATATACTTCACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAGTTCTTTGATTGTGATGAAATGATTTACATCAGCGAC
TTTAACTATTTCTTAGAAGAGGTTAGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGGGTTACCGTTGACCTCTTCATGGAAAAAC
AGGACTACTGCTTGTACCGGCACAACTTTGATTCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCTGGATCCTTAGAGAGTAATTC
AATTATTATTTATAACAAAAGTCATCAATATCGTTTTCATAGCTATATAGAAAGAAACGAAAATAACTCCTTATTGGTGTGCAAACTGCTGACTGCATTCTGTGTGCCCTTT
GCAAGTTTTTTAATTACAATAAGCTGGTTTTTTATTTCTTATATTTTCTTTTTCAGTTCTAATTATAACAATGATATGCTTTGTCCTAACAATCGCAGTATATCTATACATT
AAGAAGCTGCGAAATGTTACTGGCAAGTGCATTGTATGCTGTATAGTTTCGAGGTTTATCCAGTGCTTGATCATGATACTAGATCATTTAAATCTATTGAATGGCATTTGCT
CTCCAGCTGGTTAGTTCGAATTGGAGATATCTCTGGAAGCTTTAAATGTTATATCGTTGACTTCTTGCAGGTTACAGCTCGCACTTTTTCCGGATGGCTTCCAACCTCTGGC
TCTCCGTCATCAGCTACCATACGTGGAAAGTCTTGACGTCGCTCAATCGAGTCGACCCTAACTATCGGTTCCTGCGGTACAACGCCTTCGTCTGGAGCACAGCCGCAATCAT
GACGGGAAGTATTTATATAGTTAATCAGATTTGGGAAAACGATCCCAGTAAATGGAACTGGTTGCCTCTGGTCGGTTTTATTCGGTGCTCGGTCAAAGGTAAACTTTCAATG
CGATTGCTTTTAAATCGTTACCATGGTTACGTATTGCTTTTCCATATCAGATTGGCACCCATCCGTCTGGATCTATATAAGTGGACCGTCGCTGGCCCTGAGCACTTTCAAT
GTCGCCATGTTCGCCCTGACAGCCATTTACATTAGGAAAGTGAAGGGGGGTATAAATAAGTTCACAAATGAGGAGGAGGGAAGGATTAACTGCATAAACTTTGACAGCCAGA
CGTAAGATATGATAAAACCTTGGACCTTGGTAAACTAGTTATCCTTGCCGTATTATCCACATAGTTACCTACAGTTCCTGCGGCTCTCCATCGTGATGGGCCTTACTTGGAT
ATTCAATGTCATTCCGTATTCTGCACGGCTCCACATTTTCTGGGAATGGGTCGGGATAATATCCGAGTATTTTCACAGCGCGTTTGGAATTGTTTTGTTCGTTCTGCTCGTC
CTGAAGCGCAGCACATGGACTCTAATGATGGATTCTTAAGCTTGTTTCTATATAATTCTAATAATTGTTTACAATTAAAGCTTTAAACAATTAAAGCCCCTCATTTTATTAA
TATCCTTTGCTGAACGAAGGAATTGAATTGAATTTTATTGAACTCAATAAATAAATATAGTTTTAGAAAACGATGTACTTTAAACTGCCAAGAGGTCACCGCGACAAGCTTT
TCACTTCGCCACCCTCTGGTAACACTGATTCTCCAGCCACTTTCGATTCTTTAAAAAACATCGATAGGTATCGCCAAAAAAATCAGCTGCTCTGCTTGTAGTTTTGCGATTT
CATTGCAATTTTAATTTGCTTTACTGTTCAAAAAGAAATACAGTGCAAAATATGCCGAACGAAAACGCAGAGGACCTCTCGGGTCAAGAGCTAAAGCAGAAGGCCAAGGAGG
TGGCCGACGCATCGGAGGCAATGCTCGAAAAGGTTGTTGCCGGCTTAAATATCCAGGACACGGCATCGACGAATGCAGCAGGAAACGAGGATGCGGAGCAGCCTGATGGTGC
CAAGAATGAGGCTTCAGTGTCTGCGAATGCAAGTAAGTTTTCAATATGCCTGTGAAAAGTCATTTAGAGGCTCCTCCTGCGGCTGAAAAAATGAGGTTAGGTTTTGTCGGGA
ACTACAGTCCAACCTCTCTAACTTGAACTGCTGTGCTATTCATTTCAATTCAATTTTCAATTTTCATTTTCTGTGCAATTACTCCAAAGATTAGCCCTAACGGAAAAAATAAA
CAACTCATAAATCATAAATACATATTTGATTTTGATGAAATATGAATATACTTACATAGATACAATTGAGTTTTTTATTACGAATAAAATATGTGTAATATAAGATTTCTAC
CTTGTTGATTCCATTGATTGATGCTTCATTGAAACCCCAAAATTGGGCAAATTTGTTGTCATGGGAAACTAGTCCACTGTCGTGATGCCTTAAGCTTTTTTTGGGGCAGATT
AGTCATTCCAGTTTAAACTAATCGAAAA
(SEQ ID NO: 85)

Exon: 1001..1673
Exon: 1849..2025
Exon: 2087..2338
Exon: 2403..2836
Start ATG: 1001

Transcript No. : CT22963
ATGAGGCTTCCTTGGGTGATTTTCTGTACAGTTCTGCTTTTAATTTTTACGAACAACTCAAATGCAGATATTCCCGGGTGCAACTACTACGACACGGTTGATATCTCATACA
TTGAAAGGCAAAACGATTCGTATTTATATGATGACATCGAAATTCCTGCTAGCTTAACTGGATACTACGAGTTCAGGCAGTTTGGGGACGGTTCGATTACGCCGATTGAAAA
GCATTTAAGGGCCTGTGTTTGCAGCGTGCGGCCCTGTATTCGAATCTGCTGCCCAGCCAAGAACTTTTTGGCCAACGGAAAATGCGATGATGGTCTCAAAGAGGAGCTCGCC
CGGTTCAAACCCTATATATACTTCACATACATGGACCTACAGGCACGAGTACCACTCACCGATATGGCTATTATCAGAGATGAGTTCTTTGATTGTGATGAAATGATTTACA
TCAGCGACTTTAACTATTTCTTAGAAGAGGTTAGCATTCAAATCTTCAATAAATGTGGACTTATAGTTTGGTTCCAGGATGGAAAATTTTGGGTTACCGTTGACCTCTTCAT
GGAAAAACAGGACTACTGCTTGTACCGGCACAACTTTGATTCGGATTTTCCAAAGTCCATGTGGATAATACGACACCGCTGCACAAGCCACATATCTCCTGGATCCTTAGAG
ATTCTAATTATAACAATGATATGCTTTGTCCTAACAATCGCAGTATATCTATACATTAAGAAGCTGCGAAATGTTACTGGCAAGTGCATTGTATGCTGTATAGTTTCGAGGT
TTATCCAGTGCTTGATCATGATACTAGATCATTTAAATCTATTGAATGGCATTTGCTCTCCAGCTGGTTACAGCTCGCACTTTTTCCGGATGGCTTCCAACCTCTGGCTCTC
CGTCATCAGCTACCATACGTGGAAAGTCTTGACGTCGCTCAATCGAGTCGACCCTAACTATCGGTTCCTGCGGTACAACGCCTTCGTCTGGAGCACAGCCGCAATCATGACG
GGAAGTATTTATATAGTTAATCAGATTTGGGAAAACGATCCCAGTAAATGGAACTGGTTGCCTCTGGTCGGTTTTATTCGGTGCTCGGTCAAAGATTGGCACCCATCCGTCT
GGATCTATATAAGTGGACCGTCGCTGGCCCTGAGCACTTTCAATGTCGCCATGTTCGCCCTGACAGCCATTTACATTAGGAAAGTGAAGGGGGGTATAAATAAGTTCACAAA
TGAGGAGGAGGGAAGGATTAACTGCATAAACTTTGACAGCCAGACGTAAGATATGATAAAACCTTGGACCTTGGTAAACTAGTTATCCTTGCCGTATTATCCACATAGTTAC
CTACAGTTCCTGCGGCTCTCCATCGTGATGGGCCTTACTTGGATATTCAATGTCATTCCGTATTCTGCACGGCTCCACATTTTCTGGGAATGGGTCGGGATAATATCCGAGT
ATTTTCACAGCGCGTTTGGAATTGTTTTGTTCGTTCTGCTCGTCCTGAAGCGCAGCACATGGACTCTAATGATGGATTCT
(SEQ ID NO: 86)

Start ATG: 1

MRLPWVIFCTVLLLIFTNNSNADIPGCNYYDTVDISYIERQNDSYLYDDIEIPASLTGYYEFRQFGDGSITPIEKHLRACVCSVRPCIRICCPAKNFLANGKCDDGLKEELA
RFKPYIYFTYMDLQARVPLTDMAIIRDEFFDCDEMIYISDFNYFLEEVSIQIFNKCGLIVWFQDGKFWVTVDLFMEKQDYCLYRHNFDSDFPKSMWIIRHRCTSHISPGSLE
ILIITMICFVLTIAVYLYIKKLRNVTGKCIVCCIVSRFIQCLIMILDHLNLLNGICSPAGYSSHFFRMASNLWLSVISYHTWKVLTSLNRVDPNYRFLRYNAFVWSTAAIMT
GSIYIVNQIWENDPSKWNWLPLVGFIRCSVKDWHPSVWIYISGPSLALSTFNVAMFALTAIYIRKVKGGINKFTNEEEGRINCINFDSQT*
(SEQ ID NO: 87)

Name: mth-like 7
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384621
TGAACCTGCACCACCAGCCAAGACCCCAGCCAAGACTCCGACCAAGGCAGCTGCTCCAGCCGCTGTTGCCAGTCCCGCAGCTGCGGCAACGTCGGCAGACGCCTCTCCATCT
CCGGCCAAGAAGGCAACGCCTGCTCGCGCTGCCGCCGGAGCCAAGGCCACGCCGCAGCGACAACGCGCCCGCGGTCGCTACAATCGCTACTAAGTAGACGTCGAAGGAATTC
GTAATTTAACATAACATAAACCTTAATATACAAGAAAAAATTTACATTCACGCCCAGGAGGAGGATGAAGAAAGAACAGATCAAATCGATCTCTTCCCGCAATTTGACATTT
GCATCCCAATTTGCATATACACTGAATGAATGGAAACTGAGAAGAAGACGGTCAACTTTTAGTTGAAACAGTTGAATATATTGTACGACGAAGCTTCTTGAAACTGACGAGG
GAGAACACAAACAAGTTAATTGAATCGAGAATCGTATTACCTACCCGATTCCCCCCTCACAGAACGCATACCATGCCAGCCAGCGAACCAACCATTCAACTTATGTAATCAT
ATAAATTCTATAATAAACAAAGAAACAAAACTAGTTGTAAAACAAACACGATTCCTGTGTGTCATTGCGGGATATGGAAATTGGCTTTCAATCTGTTGCAAAATGAGATTCA
GTAGCAGTGTATATATCTATGAGGGCTGTACGATTTCGTTTTAGCGCACAATTAATACATTTAATACTTGATCTTTTTACACAAGCGTTATATTTAATGCTACAAATCGGTA
CAGAGATTAAGTGAGGTGAGAATACATACATATAAGGAAAGCCTTGATATAGTAAGATACGAAGTGGGATATATTTATTATTTACAGCTTACGAGACTAGGCGCTACTCCCA
GGAATGCTCCGAGGGAGGCGGAAACGCTTTTTTGAGTGAGATCCCACTGAGTCGGGAAATGGGATTAAGTACTTGGGGGCGCGCGGTATGAAGATCGAGTGGAATCAGTTGA
ATTCCAGAGTATTCGCCTCCGTCATCGTCGTACGCATGCGGCTCTGGTCGCTCTGGGAACGACGAAGTCCGCCCGGACGCCATCCGCTCCTCCAGCGCTTAATGCATCCCAG
CAAGGACCAGTTGATGGACTTGGAGCGACTCAGCACATAGACATACGGATCCGAGGTGAAGTGCAGCGCCGTAAGAACATCGGCGATGATGAAGAACTTGTTCGATGCGGGC
ACCCGATTTGGAGCTATAGCCAACGGGATGGCGATCTAAAACGGAGATAATTGTTGATAAAGGCCATGTAAATTTATTATATTATATTAGATTTGCATAATACGGAATGGTT
GAAGATACTAGACGTTTAGTAGTTACTATCTGCAAATTTAGAATGGTATCCACCCAGTACTCACCATCTGTGGCATCCAGCAGATGACGAACGAGATGCTGAGGAACGCCAT
TAGTTTGGCGAACTTGATCTCGACTGGCGAGGAGTCGGTGGCCGCCATTGTGACGCTAACGCTGTGCCTGTACTGCCTTGCCGGCTGTACACTGCGATGGCTGTTTCCGCTT
CCGGTGCTCAGCTGCGTCTGGTAGAGGGTGGTGCCGCTGCTGCTCTCGGGATCGATGCTGATGGCACTGTTTTTGTCCCTGGAAACCAGGTCGTAGTGCATATGCCGCTTGG

CCGTGCGACTCCTTCCGATCACACAGAGCAAGGTGTGGGCCACGAAGAGGTTGCAGGCCACGATTACAATGCACAGCAGGGTACCTGGAATACATCGATGGTCAAACCAATA
TTTGAAATGTAACCAATTTGATAGGAACTTACCGAAGACCATAAAGAGCACCGCATACGTCTTGTTCCACACGCCCGGCGCATCGCGATATCGTATACACTTAAGCTGGTCC
GGATTGGATTCATCGATGTATGCACCAAAACCAACGAATGGCAGGAACGTGATCACCACGGCGATCATCAGAATGCTGTTGATGCTCTTGCGAATGAGCTCGTAGGTAATGT
GCTATGGCGAGAGAGATTTGTGAATATTTAATGAATGAGTTGGCAATGTTTGTGATTTCCTGTGTGCACTTGTATCATTTCTTAAAATTAAATTACACTAAGAATAAAATGC
ATAGAAAATGTTATTTATTATACATATATTGCAATCAAGTTCATACCTTGTGGTAGATAAAGGGTCTGGCTAGGGCCATCCATCTTTCTGCCGCCATAACTGCTGCAATGCA
TCCAGAACTGAGCCCGAAAAATCGCCAAACCACGAGTCCCACGCAATCCACTCGGATGAACGACTGCAGGACTTCCTTTGAGAGATACATTTTCAGCAGCGTCGTGGTCAGC
ATGCCCAGTAGAGCCACCAGATTGTTGGTCGCCAGGCATCTGAAATAGTTGAGAGCTCCGTAAATTTCGGATAGCTTAGCTTTATGGCATTTCGAAGCTCATAAATCGGAAA
GTAAATGCTTTTCCTTGCCATTAGCGCCCAAATTGCCTACAGTTCGTTGGGTTCTTTAATAAGCTACGAATATAAGTTGGGTATAAGATCGATAAAATGTAATAACGAATTG
CTTTTAATTAATTGCCACTTTTATAAACCGAAACTTTCAATCAAGATACCTCAAAGCTGCTCCATATGGTTGTACAACTATCGATTTTGGTTATCTAAGACGTGACTCGGTC
CTTAAACGCTTTGATTCGAACGCCGACCAATAGCAATGAATAATTATTTTGGTTCGCAAAATGCTCAATTATAAAAAGGAAATATCTATTATATGGATATTTAAAAATTTAA
TTCATTTTTTTGAATGTTTATATTATAAAGGATGCTCCTCAAATTTCTCCATTGAATATTGTGTATATTCACTTTATATTCGTTCATATATTTAACATTTGTATATTTATAC
TAATATAAACTAAAGTACAGATCCCTGTATTTTTATCTACTTAGTGAAATATGTGACTGTCAATGGTGACTTAGATTAAGTGTTTCCTTTTATTGGTATCATAAAATATTGC
ATTTGTGGGGCACGAATGTGTTTCGATTTGATTGTTTTTGGGGCGCCGTTGACACAGGAAACTCATTTGTCTCCGTGTGAAATGAAACTATATGAATGTAGGGCTTTTTTCC
AATGGAAATATATGTATATATATATAAAACACACGCTTGGAATAGAACACAAAAATGGGATTGACAATTCCCGGGGATTGGCCTAACAGTCCGTGCATCGTTAGGCATCGCA
ATGACTTTATTCATTTTTGTTTTCATTATTTGTTTGTGCACATTGCACCTTTATACCTGATTTCAGCGAACAAGCGATTAAAATGCATTTTATGGTGCATTGTAAAAGTTTC
GCCAAAAGCTAAAATGTCGGAACTGTTCATTACAGGGCTAAAATGTCTAAATCTTCAATTCGACGATTTGCAGTGCGAAATGCAAAATCGTTTTGCTAACTTGTAAAGGCAA
AATCATTGTTTACTCTAGCAACGGGAAAATACACACAGCTATTCCGCTTTTCGTTCATAGTCAAAAAGCATGAAGGGATACCCTGATTTATCGACACCGCTCAAGTGCCATT
CACCAGCTTCATTCAATCAATCAAATTACAGATCAAAGGCTCTGCATGAAGGAGGCCTGATGCCTGAAGTGGAATTCGCGTTAATTTATGCTTCCTTGACGGCTCAAATTGC
TGGCGACAGCAATTTACGCATATTACGCATACGCACTGTCGTCTCAGCAAAAAAAAAATGTTGGGAAAGGAGCAAATCGCAAACAAGCTCACTGAATTCGCCATGGGCCAAT
GCTGATAAATGTAATATAAATTGTGACAAATATTTGGCTTATGAACTAAAAGAATTTCCGTGTGTGTCATAGATTTTATCACTTCTTTACAGTACACACCCTACCGAATGAT
TAACACGATGTTTAAAATATCAACAAAGTCTATTATGAGTATTCGTATTGAGAAGCGAATATTTTACAAGAAATTATTATTATTATATTAATATTAAAAGGAATTTGGTGAA
ATACACAAGTTTTATTTCACTCATATTTACAAAGGAATAGATTATCATTTTTGAGTATTATTATATCTACCAGTTTTTTGCACCTAAGCAAAGTTAATTTAAATACATATTC
ATATAAATCTTTTCCCCAGTATGCACATTTGCCAATCAAGTCAACCACCGATCGCCGTCTTCAGTTACACCATAAAGCTTTTAATGCATTCGCAATTGGCTGCATTCCCACG
TCGCCGATAAGATTGACAGAAAACAGACTGGAATATGGAATATGGACAAAATCTCAATGCCCAAAACACATGGCCAACGAACTAATTCCGTCATTTGATAAGATATGCAAAA
ATAAAACAGCAAAAACCATGCCAAGCAAAACAATTCAAACTATGCAAATCAGACCCGCATAGACATATTATAGAGCACCAAAGTGGCCGCAGTTGAGAGCAGGTTGAATAAA
CTTACATAGCCATGAACTGGACACAAACAGTGGCAAACATTAACTGCATGAAATGAGTAGCTACACGCTGAGAAAATATGCAAAATGTCAGCTCAATAATATTATTAACAAT
AATACTATATGGTTTTGCATGCTTAATATTTATGCAAATCAAATAGAATTTGGTGAACTACACAACTATAAATTACTACTAAGATGGGTACATTTTGAAAGATTTTAGGGCA
AGCAAAACCTACTTTTTGCATTAATATATTATTATTTCCACTATGCATAGCTGCACTTACCGCAGCATGAGCGTGTATTTGCTGTTTTTGTTGAGCTTCTTGCGCGCCAAAA
TGAAAAGGGCCAGCGAATTGCCAAAAACGCCCAGAACCATGATAATAACACCAATAATCAGGCGATTTCGATTTGCATAGAGCGCTTTCGGAGGCGCCACAGTAGTTGAGTT
GTGCGGCATCAACAGGTCGAGCACAGGTGTTGTGGTTTCCATTTTGCCGGTATGATAATTGCGTTCGATTGCAGGTGATTCAGGTGAAAATCTAACTGGCCGCAAACGATAT
TTCTTTTTCATAAATGAGCCTGCCGCAAAGGCAAGGCAGCTCGCAAATGCACTGGTAGATGGAAACCCGTTTCGATAGCCAAATTGGGACTTTTTTCCTAAGACACTTAATT
AATGATTAGCGCACGCGACACTTGATCAACCAACGACCGACCAATTCGAAGGCGATTCGATTCGATCCGAGGAGCGCGTCCGCAGTCAGCCAGGCGTCGATTTAGAATGACT
GCCGACCGTGCCGAGCAGAGCACTCGTCGGCTGCTGCCCATAGCTCCAAGCTATCCAAGCCATCATATACACGTGGGGGTGGTGGTAGGGGCCAGTTCGGAGCGAGGCCAGT
TCGGCTGGCGCCGCCGCCGGTCGGATCGCATCGGATGTAGCATACAGCACATAGACACCCCCAAGCGAAGCGGCCACCCACATCCACGCCCAGCGATGGCGACCACCCCATC
GAAAATTCACAGCTCGTGCCAGCGCCTTTGCCTTTGGACGTGTGCGAACGGTCGGAATAGCACGACGAATTTTATTGTTTTATACTCTCTCCACCTGCGGCTCACTTTAAT
TTCACTTTCAACTGGATGTAATATTAATTATGCAGTTTTTAAGCTTCCATTTGATTTTATACAGTTGACTAAATTAGAATAAAATAATATACTCAAAAGTTTACTTCAGCAA
TATATTTACTCATGTTTCTTTGTCATTGAAAAATTAACAAAACAAAAAAACAGCTATTTCATAGATAGATTAATTATTATTATTAAGAGTCTTCAGATTCGTCATCTTTAGT
ATATGGTGTGTCATTTGAAAACCTATTTTTGAGGAATATATGAACTGAAATCTGCGTATACTTTAATTACTGTAATGTATTAAAGTATTGTTTCTTTTATACAAAAAGTCTA
AAAATATAAGCAAACAAATTATTTACTGAAAATTTGTTTTTATGCCACTTTTAAGAAATTTTGCGAATTAAACATATATTTTTAAAGCTAGGAGCCTGGTAATTATAGTA
GTTTAGCTTTAACTCTTTGCTGAAAATATTTCCAACTAATTTATTTGTTTTCAACAGAATGCATTACGAAAGATAAAAATTAAATTTGCAGTTACTAAGCGTTGCAGTTGCA
CCTTTTCTTATTCTTATCTTCTTAGTATATAGTACTGTTTTGTTGGGCAAAAATGTTTATTGTTTAGTTTCAACTGCC
(SEQ ID NO: 88)

Exon: 5238..4765
Exon: 2391..2175
Exon: 2017..1825
Exon: 1764..1409
Exon: 1267..1001
Start ATG: 4970 (Reverse strand: CAT)

Transcript No. : CT23019
ACTGCGGACGCGCTCCTCGGATCGAATCGAATCGCCTTCGAATTGGTCGGTCGTTGGTTGATCAAGTGTCGCGTGCGCTAATCATTAATTAAGTGTCTTAGGAAAAAAGTCC
CAATTTGGCTATCGAAACGGGTTTCCATCTACCAGTGCATTTGCGAGCTGCCTTGCCTTTGCGGCAGGCTCATTTATGAAAAAGAAATATCGTTTGCGGCCAGTTAGATTTT
CACCTGAATCACCTGCAATCGAACGCAATTATCATACCGGCAAAATGGAAACCACAACACCTGTGCTCGACCTGTTGATGCCGCACAACTCAACTACTGTGGCGCCTCCGAA
AGCGCTCTATGCAAATCGAAATCGCCTGATTATTGGTGTTATTATCATGGTTCTGGGCGTTTTTGGCAATTCGCTGGCCCTTTTCATTTTGGCGCGCAAGAAGCTCAACAAA
AACAGCAAATACACGCTCATGCTGCGATGCCTGGCGACCAACAATCTGGTGGCTCTACTGGGCATGCTGACCACGACGCTGCTGAAAATGTATCTCTCAAAGGAAGTCCTGC
AGTCGTTCATCCGAGTGGATTGCGTGGGACTCGTGGTTTGGCGATTTTTCGGGCTCAGTTCTGGATGCATTGCAGCAGTTATGGCGGCAGAAAGATGGATGGCCCTAGCCAG
ACCCTTTATCTACCACAAGCACATTACCTACGAGCTCATTCGCAAGAGCATCAACAGCATTCTGATGATCGCCGTGGTGATCACGTTCCTGCCATTCGTTGGTTTTGGTGCA
TACATCGATGAATCCAATCCGGACCAGCTTAAGTGTATACGATATCGCGATGCGCCGGGCGTGTGGAACAAGACGTATGCGGTGCTCTTTATGGTCTTCGGTACCCTGCTGT
GCATTGTAATCGTGGCCTGCAACCTCTTCGTGGCCCACACCTTGCTCTGTGTGATCGGAAGGAGTCGCACGGCCAAGCGGCATATGCACTACGACCTGGTTTCCAGGGACAA
AAACAGTGCCATCAGCATCGATCCCGAGAGCAGCAGCGGCACCACCCTCTACCAGACGCAGCTGAGCACCGGAAGCGGAAACAGCCATCGCAGTGTACAGCCGGCAAGGCAG
TACAGGCACAGCGTTAGCGTCACAATGGCGGCCACCGACTCCTCGCCAGTCGAGATCAAGTTCGCCAAACTAATGGCGTTCCTCAGCATCTCGTTCGTCATCTGCTGGATGC
CACAGATGATCGCCATCCCGTTGGCTATAGCTCCAAATCGGGTGCCCGCATCGAACAAGTTCTTCATCATCGCCGATGTTCTTACGGCGCTGCACTTCACCTCGGATCCGTA
TGTCTATGTGCTGAGTCGCTCCAAGTCCATCAACTGGTCCTTGCTGGGATGCATTAAGCGCTGGAGGAGCGGATGGCGTCCGGGCGGACTTCGTCGTTCCCAGAGCGACCAG
AGCCGCATGCGTACGACGATGACGGAGGCGAATACTCTGGAATTCAACTGA
(SEQ ID NO: 89)

Start ATG: 269 (Reverse strand: CAT)

METTTPVLDLLMPHNSTTVAPPKALYANRNRLIIGVIIMVLGVFGNSLALFILARKKLNKNSKYTLMLRCLATNNLVALLGMLTTTLLKMYLSKEVLQSFIRVDCVGLVVWR
FFGLSSGCIAAVMAAERWMALARPFIYHKHITYELIRKSINSILMIAVVITFLPFVGFGAYIDESNPDQLKCIRYRDAPGVWNKTYAVLFMVFGTLLCIVIVACNLFVAHTL
LCVIGRSRTAKRHMHYDLVSRDKNSAISIDPESSSGTTLYQTQLSTGSGNSHRSVQPARQYRHSVSVTMAATDSSPVEIKFAKLMAFLSISFVICWMPQMIAIPLAIAPNRV
PASNKFFIIADVLTALHFTSDPYVYVLSRSKSINWSLLGCIKRWRSGWRPGGLRRSQSDQSRMRTTMTEANTLEFN*
(SEQ ID NO: 90)

Name: Prostaglandin receptor-like

Classification: G_protein_linked_receptor

Celera Sequence No. : 142000012711419
TACACTCATGAATATTCAAACAAAACTCTTCAACTTGTCTGGCATATATCCATAGACAATTACGCGGGTATGTATGCTGGGTTTTGTGTTCTTAAAATCGTATATCTGCTGC
AGGATTGACATCGAACTGCTTAACCCCACGACGGCAACCATTCGTCGAAAAAGTTTGAACACAGTTTATATTTTTGCAATATTCTAAGTACTATAACCTACCAAAATAACTG
TAAAAGAAAGTTTCTATGAACGTTTGAAATCAAAATTAGGTCCGATTTTGCGTAGTTTTCGACAAAGTGTCACCGTCAAAAGGGCTTCGACTGGCATATCCCTCAAAGGGTT
AACATGCTTTTCTTTGGTCTTCTTATGGCCTACCCCCTTCATATGTCTGCTTTTCATTGATATATTCTGCGGAACAACAAAAGAAGTATTATTCTATATTGGGATATGCTTT
AGAGTTTTGGTTTGTAGGGTTTTCTAGTCTCGGTAAATGTGTATATCTTATGCTTAATTTATTTTATTAGCTATCATATCATATACTTTGCATGGTTATCATAGATTGGTTA
TTAACATTAAAATATTTATTGTCTAAAACATTTCGCCCCGCTTAGTGCCACTCTAGACCTATGTATGCATCCCAAATCGGATTGCTTGTTTTCCCCCCCAACGATTCTAACT
ATAGCACACACCTTTGGCTACTTACAGATAACTCTTCCCTGTCCACCGATTAAATTTTGAATTTAATAAGTGTTTGATTCTAATACTTTGCTGATGGTCGAAATCCAGCAGA
TATTCCATTAAGCGGAGCACTTTCGTTGGGTATGTATAGAAAACTTATGGCCTAAGGAGGTGTAAGGGGTTTGGGAATGGAAAAGTTTGCAAGGCGCCAGACTCTCGGGGC
TCTGAAGTGTCTGGTCCTTTGCGCAAGTGCAGTGCAAAATGGAAATCTATCAGGGATTTAACTCAACCACCAGCAGCTTAACCACTTCAGATCCCCAAACCATCTTAGGTCA
TGTGCAGATCCCCCTTGAGGATCCGCGTAAAGGTCTTCTTGAACTGCTGATTGGCCAGAGCGTAGCAAAAAGGATTCATCGGGCTGTTGGCGTAGCAGAGGAAGTACGAGAA
CATGTAGAGGTGTTCATTGATGCAGGGCGGATGCCGGCAGAATCCCTCGACGAGGGCCAACACATGGTAGGGTGTCCAGCAGGCCACGAAGCAGCCCAAAATGAACGAGATC
GTGCGAAACGCTTTGCGAGCACGATTCTCCGACTTGGACTTCTGCCGTCCACCCACTCCAAGGATCAGGGGTAGAGCCTTCTTGCTCTTGAAGTGCTTTCCTATCGAATGAA
TGAATGCCCTTTTGCTGGAGTTTCGGCTGGCCGCCAAACTCTTCTCCTCCCTGTTGGCATCGTCATCCCCCGCGTTCTTTTTGGTGGCACCAACACCGTCTTGTTCCCCGCC
TCCACCCACACCGCCCGTACTGCTGGTGGTGCCCGAGGTGGAGGGCATCTTGTTGGTATTACTGTCCGAGGTTTCTCGCTCCTCAATCCTTTCCAGTGGCACCGGTGGTGTT
ACACCCGCCGCCGACTGGGGCGTTCCCGTTTGTGGCGTTGAAGTCTTTACCTTGGGTGAGAGTGGCATTGGTGGCATCGGGGAATAGAGTATGGCATCCTTGCTGCTGGTCT
CCATAGGACTGGAGGCCCGGCGATGGGGACAGTCCGCAGCTAGCTTTCCGGGCACATCCAACAGGCAGATGGCCTCAATGCAGCTCTCCCGCCTCACCTTAGCAAAGTCCTC
GATTCCTGTGTACGAGACGAGAAGACGTTTTTCAATGGCCACCGCCTCCTCGATACTCTGACTCTGACTCTGGTTGTGATTCGGATTCAGGTTGAGGCTCAGGCTTTGGCTG
GGACTCTGGTTTTGGCTGGGAGTTTGACTGAGGCTCTGGTTCAGGTTCGATTTTGGTGGATTTCGGCGTGCGGGCGGTGGTGGAGGCAGGACTGGTGGTTTCCCCTGGACGC
TGTCATTCGGGGAGGTCGTCGAGTTGGCCGTCACACTGGCCAGCATGGCGGAGCTCTCGTCCATATAGCGTAGCTCCGAGTTCGTCATCAGGTCATAGGTCAGGGCCAGTTC
ACTGTTACCGGAACTATTCCCAAAGGGACGATCCGAGTAGTCATCGGAGAGCGGGGTGGGGCACTGAAAGGCGTCGGGAGGTGGCAAAATGGTCTGCACCAGACCCTGATGC
ACTTGATCGCTAGGGACCTCGATGGGTGCCAAGTCCACGGGGGACAGGGGATCACTGGTGTTGAGTAATGTTGGAGGAGCGGACAGGTCGACCAAACGATCATGCTCGGCTA
TCTGGGACAGTGTGGAGCAGGATGTGCGCTTGGGCCTGTCCAGCGGTGGGTTCGGCTGGTTTTGGTGGTGCTGGTGGTGCGGTTGTTGCTGGTTCTGGTGGTGGTGCAGGGA
GCGCATCTCTTCCTTGCTCTGCGCCCGCTGTAACGGACTCTGGTTGTGATGGTGCTGTTGGTGGTGGTGGTACTGGTGCATGGCCTCGATCGACTTGGAGTTAGTGGTGCTC
TTCTGCAAGTTACCTTTGCCACGAGTGGCCAGCAGGGCGGCCTGTGCCCCAAAGACCAGACCAATCGATGACCGCTTCCTCATGTTGTTCAATTTCTGCTGGGTGATCAACT
GCTGGGCCTGGTTGACGGATGATTCCTCATCGGAATCGAAGGCAGGACTGCTGGATCTCTCGCTCTTTTCCGCATCCACGGCTGCCTCACGTTTTTTCGCCTCCTCCTGGAT
TTTTGCTGCCGACGCTCTCCTCTGTTCCGGTGTCATCTTCTCCACTTCCTCGGTGGCCACGGCCAGAGGATTTGCCTGACCGGAACCACTTAATCTCTTAATCACTGTGGTG
CAGGCACTATCCAAATCGGTTTGATCTCCCGCAAGTTCCACTTTGGTCTTTAGGATTTTCTCCTCAATCACACCAATGCCAGCTGCATGACCTGCCATTCCGGACATCGCTC
CAGCACTCAGTGCCACCATGGATTGCATTTTCCGTTGCTTCGCCTCACTGCGCTTTTGCATATCATAGGCTGTTTTGTAGATCCCGGCATATAAGACAAAAAGCACAATTAG
TGTGGTCCAATAGTAGCCAATGATTAGGGCCGTATTAAAGATCGGATCCTTGAGAAACTGAACGGCGCACTGGCCAGGAAGTAGATCCCGCTTGCCGGTGAAGTGCTCCCAT
CCAAAAATCGAGATGAAGAAGAGCAGAGCGGGAATGATCCAAGTGATTGTGACCATATATATCACTCTGGTTCTCGTCCTCCAACTCCGGTACTTGGCGGCAATCTTCACAG
AGCAGTAGCGATCTATCGTGATCAACAGCACCGTGTACTGGGACACCAGACACACAGTATAGTCGACGGAAAGCCAGAGATCGCAGAGCATCGGACCCAGGTCCCAATAACC
CTTGAGCACGTATATCGTGTAGAAGGGCATCGAAACGGTTCCTGAAATATGGGTTTATAGTACAATTAATAATTTTTTTTTTTTGAATTTAAGGTAAAATAGTTACTACCTA
TTAGCATATCCGTGGCTGCTAGGGAGGCAATAAAGTAGTTGCTGGGCTGCCGAATATTCCTGTCCACGATGAAGGCCAGTAGCACCAGGATATTACCGCCCACCGTAAGAAT
GATACAGATCGCCAGGCATATGGCTATCAGGATCGTCTGCCACAGTGCAAATGGTGGCAAAACAGGACCAAGAATCTCCGAGTCGGAACTGGATCCCGACATGGTGGAACCG
CCATCAGAGGATTGATTTCTAACAGCCAACAGAAGGTTCGTATTGTTCACATCGAAGTTCCTGTACCATTGCAGCTGTTCATTAACCGGCCAGGAATTGAAGTCACTGCTCT
CGCCATAGATCTTGGCCGCCTCCAACAGACGTTTGCACTCCGTGTGCGTCAGGACTCGCTTGAAAATGAAGCTGCACAGCTTATCCCGCTCCCTTCTAAGGCTCCTGGAAGT
GCGAAAGTAAATATATGCTTAATTAGAGTTGACATAAAACTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTAAACTAATGTCCTTGAAAACAAATTCTAA
CAATGTATCTTTTTAAGATCTCCATTTCTAAAAGCTTTGCCACTGAAATTACTATTCTAATAGGCGCCTTTATTCTATTAAGACAACTCACCAGTCGGTGATCCACTTGTT
CTTGATGGTATAGATGGAGTCGCTGGAGTTGCCATTCAGGCCAGTCAGCTCCTCCTCGCCCGCCGACTTGTTAAGCGACCGCATCCTTAACCGCTCCGCCCTCTTCCCGCTC
TGGCCACGTCCCGCTCCCTGCTCCCCGCACTTCGCCTCTGCGCTCCACAACCTCTGCGGTTTTTATGGTAATTTCAAGTTGTTCCTCCTTAATGTGCTGTCGACGGCGGTTG
TCTCGGAATGCTTACATTGCCTCCATGGAATGCCAAGACTTGTCTCCGGTCAGGTGCAGCATCTGCAAGAGGCAACAAAAGCAGCCACAGTCTTGGTTAGCACTTGAAAAAA
TATAGAGGTTATTAATATTATATGATATAATATATATATGCATATAAAATATGATATAATATATATATGCATATAAAATATGATATAATTTATATATGTATACATAATATGA
TAAAAAGTCTGAATTTCAGAATTCTGCCAGAAGTCGTTTTCCATCCAAAAATTAATTGTTTTTAATCGAATTTTTGGTAATAATAATCCGAATAAGGAATGGAATATTTCG
TTGAGATCGTAAATAAATTTTAAATCCGTGTGTGTTTGATTTGGAACCAGTCGTTTTCCATCCAAAAATTAATTGTTTTTAATCGAATTTTTGGTAATAATAATCCGAATA
AGGAATGGAATATCTCGTTGAGATCGTAAATAAATTTTAAAATATATGTGTGTTTGATTTGGAACCATTTATTTTTTTCACGATGTTGCCACTTTCCAAAAATTCGAAAATT
GACAAAAAAAATATTTTTCCGTTTAGTTGCGCGACCATTTTTAGCCAAATAATAGTAAAAAATCTATTTGAAAAATACATTGTCTTCATTAGATGGTGGAAAATAATGAAA
ATATGAAAATAATGGAAATTAAATTAAATCAAAATTAGTAATGAAAATCAAGTATTTACTAAATAAAGAAACTTTCTTTTCGTGTAGTCTTCGTGGACGTCTTCGCTTCCGC
CAAGTTTAAACTAATTTTGCAACCGTTAACTGGCAACCGGCGACTGGCGATGCAATTTGCTTGTGGCAGTTGCTTC
(SEQ ID NO: 91)

Exon: 4340..4237
Exon: 4024..3582
Exon: 3513..1001
Start ATG: 4340 (Reverse strand: CAT)

Transcript No. : CT23924
ATGCGGTCGCTTAACAAGTCGGCGGGCGAGGAGGAGCTGACTGGCCTGAATGGCAACTCCAGCGACTCCATCTATACCATCAAGAACAAGTGGATCACCGACTGGAGCCTTA
GAAGGGAGCGGGATAAGCTGTGCAGCTTCATTTTCAAGCGAGTCCTGACGCACACGGAGTGCAAACGTCTGTTGGAGGCGGCCAAGATCTATGGCGAGAGCAGTGACTTCAA
TTCCTGGCCGGTTAATGAACAGCTGCAATGGTACAGGAACTTCGATGTGAACAATACGAACCTTCTGTTGGCTGTTAGAAATCAATCCTCTGATGGCGGTTCCACCATGTCG
GGATCCAGTTCCGACTCGGAGATTCTTGGTCCTGTTTTGCCACCATTTGCACTGTGGCAGACGATCCTGATAGCCATATGCCTGGCGATCTGTATCATTCTTACGGTGGGCG
GTAATATCCTGGTGCTACTGGCCTTCATCGTGGACAGGAATATTCGGCAGCCCAGCAACTACTTTATTGCCTCCCTAGCAGCCACGGATATGCTAATAGGAACCGTTTCGAT
GCCCTTCTACACGATATACGTGCTCAAGGGTTATTGGGACCTGGGTCCGATGCTCTGCGATCTCTGGCTTTCCGTCGACTATACTGTGTGTCTGGTGTCCCAGTACACGGTG
CTGTTGATCACGATAGATCGCTACTGCTCTGTGAAGATTGCCGCCAAGTACCGGAGTTGGAGGACGAGAACCAGAGTGATATATATGGTCACAATCACTTGGATCATTCCCG
CTCTGCTCTTCTTCATCTCGATTTTTGGATGGGAGCACTTCATCACCGGCAAGCGGGATCTACTTCCTGGCCAGTGCGCCGTTCAGTTTCTCAAGGATCCGATCTTTAATACGGC
CCTAATCATTGGCTACTATTGGACCACACTAATTGTGCTTTTTGTCTTATATGCCGGGATCTACAAAACAGCCTATGATATGCAAAAGCGCAGTGAGGCGAAGCAACGGAAA
ATGCAATCCATGGTGGCACTGAGTGCTGGAGCGATGTCCGGAATGGCAGGTCATGCAGCTGGCATTGGTGTGATTGAGGAGAAAATCCTAAAGACCAAAGTGGAACTTGCGG
GAGATCAAACCGATTTGGATAGTGCCTGCACCACAGTGATTAAGAGATTAAGTGGTTCCGGTCAGGCAAATCCTCTGGCCGTGGCCACCGAGGAAGTGGAGAAGATGACACC
GGAACAGAGGAGAGCGTCGGCAGCAAAAATCCAGGAGGAGGCGAAAAAACGTGAGGCAGCCGTGGATGCGGAAAAGAGCGAGAGATCCAGCAGTCCTGCCTTCGATTCCGAT
GAGGAATCATCCGTCAACCAGGCCCAGCAGTTGATCACCCAGCAGAAATTGAACAACATGAGGAAGCGGTCATCGATTGGTCTGGTCTTTGGGGCACAGGCCGCCCTGCTGG
CCACTCGTGGCAAAGGTAACTTGCAGAAGAGCACCACTAACTCCAAGTCGATCGAGGCCATGCACCAGTACCACCACCACCAACAGCACCATCACAACCAGAGTCCGTTACA
GCGGGCGCAGAGCAAGGAAGAGATGCGCTCCCTGCACCACCACCAGAACCAGCAACAACCGCACCACCAGCACCACCAAAACCAGCCGAACCCACCGCTGGACAGGCCCAAG
CGCACATCCTGCTCCACACTGTCCCAGATAGCCGAGCATGATCGTTTGGTCGACCTGTCCGCTCCTCCAACATTACTCAACACCAGTGATCCCCTGTCCCCCGTGGACTTGG

CACCCATCGAGGTCCCTAGCGATCAAGTGCATCAGGGTCTGGTGCAGACCATTTTGCCACCTCCCGACGCCTTTCAGTGCCCCACCCCGCTCTCCGATGACTACTCGGATCG
TCCCTTTGGGAATAGTTCCGGTAACAGTGAACTGGCCCTGACCTATGACCTGATGACGAACTCGGAGCTACGCTATATGGACGAGAGCTCCGCCATGCTGGCCAGTGTGACG
GCCAACTCGACGACCTCCCCGAATGACAGCGTCCAGGGGAAACCACCAGTCCTGCCTCCACCACCGCCCGCACGCCGAAATCCACCAAAATCGAACCTGAACCAGAGCCTCA
GTCAAACTCCCAGCCAAAACCAGAGTCCCAGCCAAAGCCTGAGCCTCAACCTGAATCCGAATCACAACCAGAGTCAGAGTCAGAGTATCGAGGAGGCGGTGGCCATTGAAAA
ACGTCTTCTCGTCTCGTACACAGGAATCGAGGACTTTGCTAAGGTGAGGCGGGAGAGCTGCATTGAGGCCATCTGCCTGTTGGATGTGCCCGGAAAGCTAGCTGCGGACTGT
CCCCATCGCCGGGCCTCCAGTCCTATGGAGACCAGCAGCAAGGATGCCATACTCTATTCCCCGATGCCACCAATGCCACTCTCACCCAAGGTAAAGACTTCAACGCCACAAA
CGGGAACGCCCCAGTCGGCGGCGGGTGTAACACCACCGGTGCCACTGGAAAGGATTGAGGAGCGAGAAACCTCGGACAGTAATACCAACAAGATGCCCTCCACCTCGGGCAC
CACCAGCAGTACGGGCGGTGTGGGTGGAGGCGGGGAACAAGACGGTGTTGGTGCCACCAAAAAGAACGCGGGGGATGACGATGCCAACAGGGAGGAGAAGAGTTTGGCGGCC
AGCCGAAACTCCAGCAAAAGGGCATTCATTCATTCGATAGGAAAGCACTTCAAGAGCAAGAAGGCTCTACCCCTGATCCTTGGAGTGGGTGGACGGCAGAAGTCCAAGTCGG
AGAATCGTGCTCGCAAAGCGTTTCGCACGATCTCGTTCATTTTGGGCTGCTTCGTGGCCTGCTGGACACCCTACCATGTGTTGCCCCTCGTCGAGGGATTCTGCCGGCATCC
GCCCTGCATCAATGAACACCTCTACATGTTCTCGTACTTCCTCTGCTACGCCAACAGCCCGATGAATCCTTTTTGCTACGCTCTGGCCAATCAGCAGTTCAAGAAGACCTTT
ACGCGGATCCTCAAGGGGGATCTGCACATGACCTAA
(SEQ ID NO: 92)

Start ATG: 1 (Reverse strand: CAT)

MRSLNKSAGEEELTGLNGNSSDSIYTIKNKWITDWSLRRERDKLCSFIFKRVLTHTECKRLLEEAAKIYGESSDFNSWPVNEQLQWYRNFDVNNTNLLLAVRNQSSDGGSTMS
GSSSDSEILGPVLPPFALWQTILIAICLAICIILTVGGNILVLLAFIVDRNIRQPSNYFIASLAATDMLIGTVSMPFYTIYVLKGYWDLGPMLCDLWLSVDYTVCLVSQYTV
LLITIDRYCSVKIAKKYRSWRTRTRVIYMVTITWIIPALLFFISIFGWEHFTGKRDLLPGQCAVQFLKDPIFNTALIIGYYWTTLIVLFVLYAGIYKTAYDMQKRSEAKQRK
MQSMVALSAGAMSGMAGHAAGIGVIEEKILKTKVELAGDQTDLDSACTTVIKRLSGSGQANPLAVATEEVEKMTPEQRRASAAKIQEEAKKREAAVDAEKSERSSSPAFDSD
EESSVNQAQQLITQQKLNNMRKRSSIGLVFGAQAALLATRGKGNLQKSTTNSKSIEAMHQYHHHQQHHHNQSPLQRAQSKEEMRSLHHHQNQQQPHHQHQNQPNPPLDRPK
RTSCSTLSQIAEHDRLVDLSAPPTLLNTSDPLSPVDLAPIEVPSDQVHQGLVQTILPPPDAFQCPTPLSDDYSDRPFGNSSGNSELALTYDLMTNSELRYMDESSAMLASVT
ANSTTSPNDSVQGKPPVLPPPPPARRNPPKSNLNQSLSQTPSQNQSPSQSLSLNLNPNHNQSQSQSIEEAVAIEKRLLVSYTGIEDFAKVRRESCIEAICLLDVPGKLAADC
PHRRASSPMETSSKDAILYSPMPPMPLSPKVKTSTPQTGTPQSAAGVTPPVPLERIEERETSDSNTNKMPSTSGTTSSTGGVGGGEQDGVGATKKNAGDDDANREEKSLAA
SRNSSKRAFIHSIGKHFKSKKALPLILGVGGRQKSKSENRARKAFRTISFILGCFVACWTPYHVLALVEGFCRHPPCINEHLYMFSYFLCYANSPMNPFCYALANQQFKKTF
TRILKGDLHMT*
(SEQ ID NO: 93)

Name: Muscarinic Acetylcholine receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384504
GACATACGCAGGCGGAGAATTGATTTGATTTCTGTTATTGTTGCTTATCACTATCGATTTCGATTATTCAAAGTCCCTTTTTGCCCTGCGGCGAACGGCGATTGCCCTCAGT
TTCTGAGTTCCTGCAAATCAAATGCAATTATGTCGTTCAATCGTCTGAATGCTGAACGGTAAATCGAGTAAACAAGAGTGGGCGGAGGTGGTCTGAATAATTTATGACCGCC
CATTAATTATTACTCATAGAGCTGTCCTATGACTTATCGACTTTAAGCCAATTGGAGAGGGGGGAGTAAGGCTGGCAGGGCAACCCGATCAAAACGAAAAATCAAAGTGTGG
GTATCATTGTGTATTTCGTAAATTAGAAGGCCAATTTGGTGGCAATTCCTGTTTGTGATACAATTCCTGGCCGCCATTGTACCCGTGGTACTTCTAGGAATTCATTAAATGT
ATCCAATTTGCTGTCAAGGTAATTTACTTGGTGAACAATTAAGTTCAAGTAGCTATGTTTAGTCATTTCAAAGTTATGTCTTTACAATTTATTTTATATAAATACTAAGTGC
TATTCTTAATTAAGCATCTGCACTTAAGGTTTTCTCTATAAGGAGCCTATTCAAATAGTTTTCATTAATGTTTTCATTCTTTCCCTTTCTCGATTGCAGCTCCAACTTCCGG
CGCTTTGGTGGCGTAAAATGTTTGCCACCAATATCCGTTCGGTTGCCCGTAGAAATCGTCCTGGTGGCGCCGTGAGCTCTCGTCCTTTTCCTTTCGCGAGTCCTGAATCCCCG
TCCCTCGCTCATTTCCGGTGCGCCCGTTTTTACATTTGCAATTCGAATCGGATTGAAATCGGAATCGGAATCAGAATCGGCGCGAGGACATTTCCATTCCTTCCCATTCTCTT
TTTATTTTTTTTTTTTGCGATCTGCGTTGCATATGCGAGGGCTAATTAGCATGCGGCATTTCCAGCAATCAGAGTAGAGCGCACAATAGGAATAATAACGCGGAATGGAAGA
GGATGTGTATGCCTCGCTAGGTGCCTACAACGACAGCGGTGGCGACGATTGGAGCAGCTCGGAGCACCTGGTCCTGTGGGAGGAGGATGAGACGCAGCGAACGACTGCTAAT
GCCACCAGTCGGCATAATCAACTGCATGTGGCCAGGTGGAATGCCACCGGCAATGCGACCATCAGCGCGACCTTCGAGGACGTACCCTTCGATGCGAACAACTACTGGGCCC
TATTGGCCCTCGTCCTCGTCCTGGGAACAGCGGCCGGGAACATCCTTGTCTGCCTGGCCATCGCCTGGGAGCGCCGGCTGCAGAATGTGACCAACTACTTCCTTATGTCGCT
GGCCATCACCGATCTCATGGTCGCCGTGCTGGTCATGCCGCTGGGCATCCTTACGCTGGTGAAAGGTGAGTCCTTTGTCCCATTCAAAGTCCGGAACAGAGGTATTTTTTGC
AGCCACCTTTTGTTGAGGGACTCAATGGACACTTGAGTAAGTCACTTGTGCGCCAGGATTGGCTTGTGCGGCTAATCTTAGTTGGAAAAGTTCCTTTTCCATAATTTCATAG
TTTACTCTAACGAAAATGAGTTGGTGGAAAACTTGATGAAAATTAATATTAATTACTTGGGGCAAGGGAAAAAAGAGTGAGAAGTTTTGATAGTTGTTAAATTTTTAAGGCT
TTTAAAAAGTGACGAAGAAGATGCAAGGAAACGCAAAGATGTAGGGAATATACAAAATTATTTTAAATCTTAAGATTTCTTATATTTGCAATTTAAATGCGGTAAATGCGAA
TCAATTGAATAATGTACAGGGAATTGTGAAAAAATATGGATATTTACGGATTCCTTTTGCTGCAACTCAATCCGAATGATAATCTCTCCAAGTGGAAGAGTATAATAATCAA
AGCCAGCGACAGAAACGTGGCAAAGACATTCAGGCCAGGACCTCCTGCGATGTCGCCAACACGAGGCACGAGTTGCATTTAATGAAGCTGCTGGTGGGAATGGCGTTGGCGG
TGTGCAGATTGCATTATGTCTCAGATATTTCGATCAAAGCTGGCAAACGTAAGTAAGCCCGGATAGAAAAATGGAATCGCTGGTGTCTATGGGCCGGCAGGGCGGAGCGGAG
CTCCATCGTTAGACAATTTAATAAGCGCGGCGCCTTTGTAGCCAGCCTGGCCAGGACGAACCGCTGGCGCAGACAGGGAAGCCCAGAAACAATGCCAGCAAGATGGCAACGG
AAGCTTCCAAAGCCTTTGCAGGCGCCGGAGCAACTGCAGCATAAGAATTGCATTAGGGGCGGCAAACGAGCGCTGGTGAATTTCGCTTTGACTTAGTTCGGATCTGGAAACA
TCTGCCGACACATTGCTGTCCTGCAATCTGCACTGAAAAAAAAAACAAAATATATTTATTCAATTATTTCTATACTTAAAGTAAAAGTCATCAAGTTTAAAGCTAATTGTAT
AATTTTTTTCTTTCGATTTTTTCGATTGAGCAATAAAATATGGTTCATATATGTTAATACTTACATTTTTTTTAAGCTAAGTAACTTGAAGCTAGCTACACAAATATTATAA
CATACCAGACAGGAAATGGAAAAGCTATTTCCGTGACTTTCAAAGGCGTTGCTTATTTCGTTGCTCCATTTATGTGAGTGCTTCTTCCGGTGGCAGTTAAACGGTATTTCTT
CGATTTTCGAAAACGAATGGCTGATTTGAGGACTGCATTTATCTGAGTGCTGCTTCCGCCTCCTGCTGCCGTCCTTTTTGTGCCAACCCCTGCAGATGTTGCAACACTGATG
ACAGTTTTGTTTGTGCGCTTTGGGAGCCATTACAAACACAAACACAAACACAAACACGAGCACGGGCACCAATACAAGCAGAAAGGGCGACGCACACACACGCAAACACACA
CCTGGAGCAATCCTTGGCAAAGTGGGCAATGTAGCAAGAAGGAGGACCAACAAATGGCTGCTTGTTATATCAAGTGTGAAAAATATTGCAGTGGCAGTGACAGGAGCAGTTG
TTGCTGTAGTTGCACGGGCACACGCCCGCCAACTGGAGGCAAGAAAGGACGAAATCGGGATTTGCCTGCTGTCAAGGGCTTTGCAGAGGGTCCTGCGACAGAAGCTGAAATC
TCCAGCGAACAGCAGGCATTAGATCCCAAGGATTAAGCTCTTACAATTCAAAGGCACAATGGCGCAAGTAGCAACTGGATATGAGTAAGCCGTGCAGTACATTAATTGCATC
GGTTACGTAAGTAACATGTGGCATACAATGTGGCATTGAAATGGCCTATATATTCTATTCTCTGTTGATAAGTATGCAAATTCAATTAAACGACATGCGTTGATGCCATTT
TCATGACCAATAGGCGTATGAGTAATATTAGTAGACTAAATGTTTAAAAACGCTAAAGGTGTAATATCAAAATTAATCTAACAAAATCATTCATGGCTTTGGTTGTTGACTT
TCTGGAGTGTAATAATAATACAAATTTAAATTTTCGTAAAAACAATTTGATCCGTACGTACACGCACAAATCTAACAAACTATTGAAATGGAATCAAATAGCTGAAATAATA
TTTATAAATGAATATATATTTTTAGGACTTAAGTTCACTGAAGGCTTAGCGGGAAATAAGTTTTTAAGGATGAAATATTAATGCTATTATGCTAAACTCTGTTGAGTTTCCA
GATGCATATATCCTTTGATTTTGAACTTTCAACGAATTTTCCACTTGATCAATCGACCTTTTGGAAACGAAACAAGACCCATGCAATTAAACGACATGCGTTGATGCCATTT
TAAGCGCTTGCACTTGCAGTGAATTATGATTACGCGTGGAAAACATAATTAATGCATAATTAATTTCATTTTAAAATCAGCATTAAACGCCAGTTAATTAATAACAGTAAAC
GCAATTACTGCATAGCATAGCCATCAAAAGCCATAAAAAAGACAGTAAATTGAATTACAACGTCCCGTCAATGGTCGTGAATATTTAAGTAAACGCGAAAAGCTCGCAAATG
TTTGCAACTTAATAAATTATACTTGGGACGAATAATGATAAACAAGAGAACAGCCGCAAATATAAAAAGAAAAGCCAAGCGAAAAAAGGCGCGAGGCACGAGAAGCTGGAAA
ATTAATTGAGCGCAAATTGGTGAAGCTG
(SEQ ID NO: 94)

Exon: 1001..1445
Exon: 1885..2064
Exon: 2784..3172

FIGURE SHEET 44

Start ATG: 1001

Transcript No. : CT24036
ATGGAAGAGGATGTGTATGCCTCGCTAGGTGCCTACAACGACAGCGGTGGCGACGATTGGAGCAGCTCGGAGCACCTGGTCCTGTGGGAGGAGGATGAGACGCAGCGAACGA
CTGCTAATGCCACCAGTCGGCATAATCAACTGCATGTGGCCAGGTGGAATGCCACCGGCAATGCGACCATCAGCGCGACCTTCGAGGACGTACCCTTCGATGCGAACAACTA
CTGGGCCCTATTGGCCCTCGTCCTCGTCCTGGGAACAGCGGCCGGGAACATCCTTGTCTGCCTGGCCATCGCCTGGGAGCGCCGGCTGCAGAATGTGACCAACTACTTCCT
ATGTCGCTGGCCATCACCGATCTCATGGTCGCCGTGCTGGTCATGCCGCTGGGCATCCTTACGCTGGTGAAAGGTGAGTCCTTTGTCCCATTCAAAGTCCGGAACAGAGTGG
AAGAGTATAATAATCAAAGCCAGCGACAGAAACGTGGCAAAGACATTCAGGCCAGGACCTCCTGCGATGTCGCCAACACGAGGCACGAGTTGCATTTAATGAAGCTGCTGGT
GGGAATGGCGTTGGCGGTGTGCAGATTGCATTATGTCTCAGATATTTCGATCAAAGCTGGCAAACATGTTGCAACACTGATGACAGTTTTGTTTGTGCGCTTTGGGAGCCAT
TACAAACACAAACACAAACACAAACACGAGCACGGGCACCAATACAAGCAGAAAGGGCGACGCACACACACGCAAACACACACCTGGAGCAATCCTTGGCAAAGTGGGCAAT
GTAGCAAGAAGGAGGACCAACAAATGGCTGCTTGTTATATCAAGTGTGAAAAATATTGCAGTGGCAGTGACAGGAGCAGTTGTTGCTGTAGTTGCACGGGCACACGCCCGCC
AACTGGAGGCAAGAAAGGACGAAATCGGGATTTGCCTGCTGTCAAGGGCTTTGCAGAGGGTCCTGCGACAGAAGCTGAAATCTCCAGCGAACAGCAGGCATTAGATCCCAAG
GATTAA
(SEQ ID NO: 95)

Start ATG: 1

MEEDVYASLGAYNDSGGDDWSSSEHLVLWEEDETQRTTANATSRHNQLHVARWNATGNATISATFEDVPFDANNYWALLALVLVLGTAAGNILVCLAIAWERRLQNVTNYFL
MSLAITDLMVAVLVMPLGILTLVKGESFVPFKVRNRVEEYNNQSQRQKRGKDIQARTSCDVANTRHELHLMKLLVGMALAVCRLHYVSDISIKAGKHVATLMTVLFVRFGSH
YKHKHKHKHEHGHQYKQKGRRTHTQTHTWSNPWQSGQCSKKEDQQMAACYIKCEKYCSGSDRSSCCCSCTGTRPPTGGKKGRNRDLPAVKGFAEGPATEAEISSEQQALDPK
D*
(SEQ ID NO: 96)

Name: 5-HT2 receptor-like 1
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384504
TATAAATGAATTCTGAAAAATGTATAAAATGAAATGGACTCATCATTAGGTCAATTTGTTTGGAATCAAAAATTTCCCAAAGAAATGGTTTTACATTTCAGAAAGTGACCCT
TAATTTTCGCCATCAATATTTATTAGAACATTGACCTTAAAGCAATTTGTATAGTCGTCACTTTAACAGATCAAAGTTATACTTAAAGTCGTGCAGTTGCAACTGAAACTGA
TTCATTCTACGCCTTATCGATCCATCTTCCCCCGAACTTATCCCAATTTCTGTTGGCTATCTCGTTAGGAAATTCCTTTTCCCAGCGGCCATACGGATTCATGTTCGACCAG
ATTCTGGATTCACAAATCCCGCGCCCAAATCTCAGGGCACACGCAACAACACAGTAAACACAAGTATCACTCACCGAATGTGTGGGAGATTCTCGGCCACACAAGCTACAGA
GAAACAAACCCAATTCAAACTAAAAATTTCTAAATGGAAAATAAATTAATTTTATTAGTTCTTTATAAGATACTAAGACTTATACTAGTTGGAAATCTTCTAGGCATGGTTT
GTAAATCATCCTTATATCTTTCGATATTATTCTAAGACTTTTTTCGTAGTGCACACACAATTGTATTTTTATGGGAATAGGGCGAGAGTGGTGCGCTGAGTTATGACGTCCA
ACAGCTGTGCAAAACGATGGCCACAGCATCGGGAATAGCCGCTCATATGTGGGCCAGTTGGCCTGAAATATGTTGCATAAGCTCCCCCCTCTTGGCAATATCATTTGGCAGC
TCACGAATGAACTGTGGCTCCCGGGAAAATATGGCCAAAAAAAATAAAAAAAAAAAGGCGGGGAGTATTAAATAAAAAATGGCTAGAGAACATTGCATTGTTTGCCGCCGACC
AAGCCAATCGTAAATCGTGTGTACACATAGGTAGAAAGTATTGCAGCCACTTAATACAGTCATAAGTCGGCCATGGAACATTTACAAGCTGACCAAGAAATACAATGAAGAA
TCGGTTTTTTTTTCTTTTACAGGCTACTTTCCTTTGGGCTCGGAGCACTGCCTCACCTGGATCTGCCTGGATGTACTCTTCTGCACGGCCAGCATCATGCACCTGTGCACC
ATCTCCGTGGACAGATATTTATCGCTACGATACCCAATGCGATTTGGCCGGAATAAAACACGTCGGCGGGTCACTCTCAAAATTGTTTCGTTTGGCTTCTGAGCATCGCCA
TGAGCTTGCCCTTGAGCTTGATGTACTCCAAGGTGAGTAATAATGCTATTCTTTTGTTTATATTCTATATTCCAACAATGCGTCAAAATATAGTACCCTTTTGGCATTCTAT
TTCCCTTAATCTAATAGTGAAAACCCATTAAATTCCCTAGTTAACCTAGTACTCAAAATCTAAAGTGAAAATTACAAACACTGTGCCAAGTCAATAATTAAAATCCAAAGCG
AGACAAGTGTCTGAAATATTTTGCGCATACATATGTCCTTTACTTACCCAAATTAATTCCGTGCTCGCCCTGCTATCTTTCCTTTGATTCCAGAATCACGCCTCGGTGCTGG
TGAATGGAACTTGCCAGATACCGGATCCGGTGTACAAGCTGGTCGGCTCCATTGTATGCTTCTACATTCCACTGGGCGTGATGCTGCTGACATATTGCCTGACCGTCCGACT
TTTGGCCCGGCAGCGCCAGAATCTGGGTGGTGGACAGCAGACGGCGGCGGCCACTCCCGGATGGGCCAGTGGATGGCTCGGCCAGGCACCGGCCTTGGGTAAGGATGCTAGA
CACCACACTTCTGACCCCCACTCTCACCCAGAAGTGTCCACGAAGAGCTGCTGCTTCTGCTGCAGCTTAATATGATTTTCAGTTTGCCTAAAATTTGCATTTTGTTCTCCGG
ATTGCACTTTTTGCACTGAAAACTTTACACGACACTCGACAAAAGTAGAGAGTAGGAAATGGGCGGATCAGGATCAGCAAGGGGAAAGTGCTCAAAAGTGGGGAGGGGTGAC
AACACTAGACAAACAATGCAGAGACATGTGCCGGATACTCCCCGAATCCGTGTCCCTGGGCTTTCTGGAGCCCAAAGTGCGACTAAGCCGCAGCAATGCACACAAAATCGAC
ATTGTCGTCGACATTGAACCGGATGGAATGGGGCGGGGGGGGGGGGGGGCCAACACCCCAGGACAATCACCCATTCCACCCACTTGCACACACAACCTTACAAATAGGATGA
AAATCCAGCCCATATGTGTAGATAGCTAATTTCCACGTTGCCATGAATCGAGGGCCAGCTACGTAAATGGCTTGTAATGCTCTAAGAGCTCTTCCATAGTGCTCGACTAGTT
TTAAAACCGAAGTTTAAGCACCAACTTTTCAACGAAGCCTCGTTACAGCCCCGTTGCGTAGGGTAGTTTTTAAAAACTCAAGAGCTGCGTTAACTTTTAGTTTTATTAACC
ACTTTTTGAAATTTCCATAACACTTTTCAGCGCATCTACAAAAAAGAGGAGGAATCTTTAAACATGTAGCTAATAAATAAAAAGTGAACAGTCTGTTATTTTAGATTTAAAG
AGTCTAAATTGCTTCCTGCATCCCTAATTACCCCTCTCGTTACTAGCCCAGTTAACTGGAGCTTTTGGCTAAAAACAATCCAAGATCCTTCGCTTCCATTATCCCTTCACAA
TTCAATAGACCAGATTCAATTTATTTTAAGCATAACTAAGTGGCTCCTAATGCCATTGTCTGCCCGCAGAATTTATTCAGCGCTCCCATTTCTCATCCAATTGGATTGCCTT
GAGCCCCAAAGGGGACAGAGAGTAATTTTCTGGGGGACAAGCTGTGGATCTCTGGATGTTGAAAGTCAATACAGTTGGGGAATATTTAATGGAATGCTTCTGGTGGTGGCGA
ATGGTGAATGCTGAATGGTAAATGCTGATGGTGATGGTGACTGCAGTCGAATCGCGTGCACTGCAGGGCATTTTTGCGAAAATCGTAGCGTGCATTTAGGGGCGATTCGCAG
GACGCGGTGTCCCCGATTCGTCTAAGTGGATTTGAAGCCTCAGCTGCAGGCATTAGGGTGTGTATGTATGTATGTGTGGGTGCGCATCTGTGAGCGGCCAAGTGCGACTGTG
TTAGTGCCGTGGATTTGGGCAGTTATTTTAAGCCAAATCTCTGCTGTAAAGCGTCCGCATTTCGGTCAGGCACGTAGTCGGATTTGCCGTATTGTACTTTTTACACGCAGAG
AAAAATCTTCGCATTTATTACACAAAGAATCAATGTGTCTGATTAATATAAGCAAGATATTTAATGCAAGTTAGAATAGCTTCAAGCACTATTTCATAAGGTCTGAACTGGA
CTAGGAGAAGAATTTCTCTCAGTGCCTATGTTGAAGGGGCATTTAAGTGTGCTCGTGCGAGTTTATTAGCGAATTTGTGACTTTTCGTAGCGCACTATTGGACTGAATGCAA
TTTAATGGAAACTGTATATCCGTTTAACCCCGTGTTCGTGCTCCACAGAACGACGATGCACCTGGCGGCGCCTTTTGAAACCGGGTCCAGGTAACGCCTCCTCGGTGCTGCA
CGCACACTCGGCCAATTCAACGGATACGGATCTCAGCACTTTGGACAACCACGAGCTCTGGCTGCCCGATTCCAGGTGAGCTATCGAGTTTGCGAGCCAGTGTTGCAGTGCA
AATTACCATATATATTATTCATGTTTATGTAGCATAAAGGAGCCAACGCCCACGACCATGACGGCCTTGCATCAGTTCGGTGCCGAGATGCTGAAACTATCCCGAGGCCTGG
AGTCGGTGGCCTCGTCATCCACCACCGGCTCACCCACCAAATCCGAATTTTCCCTATCGAACCACCTGCAGCTGCAGTACCCCAGCAGCCCACAGCGATATGCCGCCACCCA
CCAGCAGTCCCACCACCACAACCACCAGCAGCATCCGGCCGGAGTTTACCACCAGCAGGCATCGCCCAAAGGTCGCCACGGAACCACAGTTCTGGGTCTGTCCACCACCACG
CTGGGCATCGAGCGGGAGAGCACCAGGAATTCCCTGGCCAGCAGTCGCATGGGCGGCGAACAGAGCGATGGGTGAGTTCATAAATAGTAATATAAATAAGTGAGTCCAAATA
ATGGGGTCTAAAAATATACTATCTAAAAGTCTGGAAGCCAACCAAAATTATGGTTACAATTTTCTCAAGATGTAATATGATGCTCAAACATAATTAAATTGCTTCACCACTT
AACCAGAATCTCTGTTGTCACAAATAATATTAACACAGACAATTAAATTCTTTTAACTTTATCCCTTTTGCGGGCCACAATTTTCACTAGAATTGTGTCAAATCCCTTGAAA
AGCATAGACAATTCGGGCAGGTACCTTGGAAAAACGTAAAGTTTTCGTGGCTCTTCACCTGACAAAAGGACAAACAGCTGCGATGGGCAAACCGTTTTGCCAGAGTGATCCT
TTCTAGTCGCTCGCAGGACCTTTTGTTGTCGCTGTCTCAGCCGGTTTTCGAGTGCCAGTACTGCGCTTGGCCAAGTGTGCTGACCACACCAAAAGCCAAGAAAATGCTCACC
TCGAGTCCTTTTTGTGACTCGTTTACTTTCTCCTTTTCCATTTATTTAGTTTTCTTTTTTCGTTAGCTCATTTTTGTGGCATTGCCTTTTATTGTCCTTCCGCCTTCGACTT
GTTGTGAAATTTTTCGTAATAATTAATTTTTTTTCGCTACGTTATCCTTTTTCTTTGTGTCGGTGCAAACTGTTTTATTTCTTTTTTTTGGCCAAAAGGCAGAGGCACTGCC
TTTTTATTTGGCATTCGAGCCGCTCAAGTGATTTTGGCGTCGACTGAGGCCTTTTGCCAAAATGTGTTTTTCATTAATTTCCTAAATTTAGTCTACTGGCTTATGCTGGGAT
AAGTGATGCGACGAGAAATTTAATTAAAAAATTGAGATAATAGAAGACGAATATATTTTGAAATTACCGAGAATTTAATAAGAACTGGAAACATGGCTAAACTGGCATATTC
AAGTAAATAGGGCGCTTCATGTTTATGGGCATTCAGGTGGAACTAATATTAACTATTTAATTGCGTATGATATGAAATTCTATGGAAGAGCCTGGCAAATTCTAAAAATACT
TTTATTTTAGGCACTCAACCCAATTCGATATTGTACACGCATTTTATTTTGATTTAATTGGTTATCATATTTAGTCTGCAAAATCCATCCAATAAAAGCATGAATTTTTCAC
TGCAACTGCGAACATTGATATCTTGAATCCCCCGTTCGATTTTAAGCTTGAATAAAAGGCAGCTCAACTTGCGAAATGACTGCACATTTTATCTCGCTCATATTTTTCCCAT

TTAACCAATCCCTTATGCATGTTTATATCATATATCATTGTGACTGTGCGCCTTTCAGCCAAATGCTTGGTATTGGGTATTGGGAGCACTATGTGCCACATTGTTTGACAGT
TGAACGAATTTGCACAATCAATAAAATGGATTTGGCTTTGGCTCCGGCTTCCTATTGTCCGCAATGGATGGGATGCCACTTGGCATAAAGGACGAAAAGAAACGGGACGAGG
TCAAAGTCGAGTGGTTTTTAATTGTTGCACAATTTGCCGTCTCATGCGGCAAAGTTGACAAATGATGCTGGCGATATTCTAATTTTTGAACTTGGCTGGGAGACTTAAAATC
AGCTCAACACTTCGCCACATAAGAGTATATATTATATGTATGTGTATGCGAAGGCGGTATCCTGCGATGTCCTGATGAGCTTGGCAGCTGTTTAATTCCTTTGTGAGGGGGA
AACAATAACAAAGTTGGCATCAGCTCGGCTTTGTTTCACTTGCATAATGAGTAGCTGGCCCCAAATACCACCTCCCCACTCCCCCATTGCTTAGCCCTGTTCCAACGCATCT
CAATTTATGCGCACGTCTCGTGCAAACAAATCAAAAACGCGAAGGATGCATAATGAGCTGACATCCTTATCCTGCGTCTTGGCAGCACCGAGCATCCCGATCCCTGTCCTCC
CTGCAGCCACCACCACTTGGCCTTCTGTCTGCTAACGCCTTCATTAGCAGCTAACCACGCCCCCAACGCCACCGCCCCACACCGTCTTGGCCAAATAAAGTTGGCGCGCATT
TGCATGCAGGGCAAAACGGGCGGCAGGACAGAAAACCGCAAAGGCAGCTCGCAAATTTATGCGCTTTATTAAAATGAATGCAAAATGAGGCGGTACAGAAAAAAAATCGATG
TACAATATGTTTATATATAACTTTAAATTAAAGCTACACATATGAGCTAAGTATTTCTGCAATTAATAATATAAAATTAAAAAAAAAAAAAATCCCGCCAGTAAAGTACAGT
AAGCCGAACTAACATACCAAATATAATATAGTAAATAATAATGCAAGAATTTTTCAAGTGTACCACAAACAACGGAGCAAGAAAAAATGTTAGAAGGCGGTAAACCACAAAG
AGCAAACACAAAAAATGTTGCCTACTTCCCGGCAAAACAGGTGAAAGAAACGAGCTGACCACAAAGCGTGCTGGCTTTTGTTCAATTGCCGTTAGCGAGTGTTCACTGCATT
GTTTTAGGCACCTCACCTTCGAATAGAGTTTTGCCAATTGGCTGCGCAATTTTAATTAAAAATGTACGACGCCGAAGGGGAGAAGGTGAAAGAAAATGGCTCGGCATTGGGT
ACTAAATCAATAAATCGGTGGGGTATGACATAGACATATTTTATATATTTTCATTATTAAAACCACATGGTTCCTTTTGTTCTAGTCTGCTTTAATTTGATTCAGGAAAATC
GTTTAATAGACCACTTTTTTAAAATGAAATGAAAATAATCATTTAAGGACTTTGACTGTAGTATGTCATTATATTTATTTAATCAGTAAAGTTCTTTTCCTTTCAGGACTCT
TTCACAGCTTTCCCAACGATTGCGAGCGTATAAAAAGCGACGAAGAGCTTCGTCAGCTGTTCCTGGACGTGAGAAACGTGCCGGGCATGAGGACGCAGACGAGGATGTGGAA
ACGCCCACGTCAACGCTGCGCCGCCACAAGCGCCATAACAGTCTGCCCAAGAACGCACTCTATCCGCGACACACCACTGTACAGGAGAGCCTCGATGACGACGACGACGAGG
AGGAGGAGAGCGCGCAACAGAAGGATCAGCAGACGAAGTCGAAGATCTGCCACTCGGACACTGAGCTAGATCCTCCGCAGAAGGCGAAGCAGTGCCACATAGCCGGGCAGAA
TCATCAGTCGGACTACCTGCAGCTGCCCTCGGTTTGCACTTGTCCGTATTTTGGGGACAGACCGCTCCAGAATTGTGTGAAATCGGCGGAGGTGAAGATTATTTCGTCCGCC
TTCCGAGTGACCACCACCACGACGGCGGTGAGCAGTTCGCCCAGCGAGATGGAGCTACTCATGTGCAGTAGTGGCAATAAAAAGTCATTGACCTCCAGTGTGAGTGCCGGCA
TAACAGGTGGACCGGCAGGAGCGGCTCCAGCCACAGTTGGCTCCACCCTCAGTCCACACTCTGCTCACAATCAGGGCAGCTCGCTGACAGTTCAGAGCGATGGTAGTGGGTA
CTTGGCTGCTCCGGGAACTCCTTGTCCTGGCCGACGAAAGTTGAGCATCTCAAAGACCGCCTCGGTGGTCACGTGGGATTCAAGTCGTCATCGTCGGCGGGGCAGCAGTTTT
GGCGGAGTGCGAACATCCCTGCTGTTGACGCCCACCAAAACAGCGACCACCTCCAACTCATCCACACCGTTGAGAAGATCGGCCACCTTGAGGAGTCACCAGAATATGAACT
ACCAGGGAGCGGGCGAATGTGGTGGCAAGACCAGGACGACAACATCCTCACCCTGCATGTTGCAGCGTCAGCAGACGGTGCGATCTCATCACTCGCGGAACTCCAGCGTGAT
CTCCAGAAACTCCTCGCGCCACGGCAGGATCATCAGGCTGGAGCAGAAGGCCACCAAGGTGCTGGGAGTGGTGTTCTTTACGTTCGTAATCCTCTGGTCGCCATTCTTCGTC
CTCAATCTCCTGCCCACGGTGTGCGCGGAGTGCGAGGAACGAATTAGCCACTGGGTCTTCGACGTGGTCACCTGGCTGGGCTACGCTAGCTCCATGGTGAACCCCATCTTCT
ACACCATCTTCAACAAGGTGTTCCGGCAGGCGTTCAAGAAGGTCCTGCTGTGCCGCTATTCGAGCACGAGTGCCTGGCGACCGAGCAGATAACAGACGCCCGTTAAACCGGG
CAAACTGAAGACGAATACCCACAAGTGCGCCAGTGTGGATTTTCAGGATTCCACTGCTCCGGTGCGGGGCGTTCTGAAAAGAAGTGTGGCAGTGGCAAGTGACATATAAATA
CGGCAAAGGCCGGTCATTTCGCCTGCGGTCCACCGCACAACGTGTCGTATACGCAATCTCGAAAGGACGATAATGCCCACCCGCACTCCACACCACGCACTCCGCACTGCTC
ACACGTGTGTCCTTGGGGCTTAAATTCTTATGTGCCTCTAATTAATTTCAAATTCAATTTGCATGCCATTGAAAACACATGTCCCAGGCGTAAGAAGTGGGGCCAAGTGAAT
TAGATATTGAAGTGGAGTCGGATCCCCGCAGGATTGCCAAGGGTCGGGGATTAAACTAGCCACAGACAACAGCATAACGAGCATGCTACTCACCCCTGGTGGTTCCTTCCAT
TTCCCTCAGCTCTCAGCAAACGGTGAAGATGCACTTAATTGCGGAATGTTTGCTTTCGAGCAACGTAATACCGGCAAACACATTTACTTTATAGCAACGCTCTTCAGTTATC
GACAAAGTTCTGGCTCACCGAAAGTTCATGTCAATCTGTTGATATTTGATAGCATTTTCCGCGAAATTCCTTTTAGATTCACTACGTACTCGAAAATAAACTTGAAAATTCA
CGTAAATTTAAAGATTGTATATATAAAGTTCAACTTTTGAAATTCTTAAAAATAAAGCGCCAATGTCGCTTCGCAAACTTCGATCGATAAACGCTATTTGGTTTGGCATTTG
GCAATTATGCATTAAATTTATGGCAAATTCACTGCCCGCCAATGCACAAAGAATTTCATCGGTATCCATTAAAGAAACCGACACGTACACAAATCCACGGCAATTTGTGACT
TTTGTTGTTTGTTTGTGTGGTTGGACCGGCAGATTGCTATTTTTATTTGAAAAGTTCATTCATTTTGTTTTATTTCTGTGGCCA
(SEQ ID NO: 97)

Exon: 1001..1264
Exon: 1550..1778
Exon: 7511..8268
Start ATG: 1001

Transcript No. : CT24060
ATGAAGAATCGGTTTTTTTTTTCTTTTACAGGCTACTTTCCTTTGGGCTCGGAGCACTGCCTCACCTGGATCTGCCTGGATGTACTCTTCTGCACGGCCAGCATCATGCACC
TGTGCACCATCTCCGTGGACAGATATTTATCGCTACGATACCCAATGCGATTTGGCCGGAATAAAACACGTCGGCGGGTCACTCTCAAAATTGTTTTCGTTTGGCTTCTGAG
CATCGCCATGAGCTTGCCCTTGAGCTTGATGTACTCCAAGAATCACGCCTCGGTGCTGGTGAATGGAACTTGCCAGATACCGGATCCGGTGTACAAGCTGGTCGGCTCCATT
GTATGCTTCTACATTCCACTGGGCGTGATGCTGCTGACATATTGCCTGACCGTCCGACTTTTGGCCCGGCAGCGCCAGAATCTGGGTGGTGGACAGCAGACGGCGGCGGCCA
CTCCCGGATGGGCCAGTGGATGGCTCGGCCAGGCACCGGCCTTGGGTGGACCGGCAGGAGCGGCTCCAGCCACAGTTGGCTCCACCCTCAGTCCACACTCTGCTCACAATCA
GGGCAGCTCGCTGACAGTTCAGAGCGATGGTAGTGGGTACTTGGCTGCTCCGGGAACTCCTTGTCCTGGCCGACGAAAGTTGAGCATCTCAAAGACCGCCTCGGTGGTCACG
TGGGATTCAAGTCGTCATCGTCGGCGGGGCAGCAGTTTTGGCGGAGTGCGAACATCCCTGCTGTTGACGCCCACCAAAACAGCGACCACCTCCAACTCATCCACACCGTTGA
GAAGATCGGCCACCTTGAGGAGTCACCAGAATATGAACTACCAGGGAGCGGGCGAATGTGGTGGCAAGACCAGGACGACAACATCCTCACCCTGCATGTTGCAGCGTCAGCA
GACGGTGCGATCTCATCACTCGCGGAACTCCAGCGTGATCTCCAGAAACTCCTCGCGCCACGGCAGGATCATCAGGCTGGAGCAGAAGGCCACCAAGGTGCTGGGAGTGGTG
TTCTTTACGTTCGTAATCCTCTGGTCGCCATTCTTCGTCCTCAATCTCCTGCCCACGGTGTGCGCGGAGTGCGAGGAACGAATTAGCCACTGGGTCTTCGACGTGGTCACCT
GGCTGGGCTACGCTAGCTCCATGGTGAACCCCATCTTCTACACCATCTTCAACAAGGTGTTCCGGCAGGCGTTCAAGAAGGTCCTGCTGTGCCGCTATTCGAGCACGAGTGC
CTGGCGACCGAGCAGATAA
(SEQ ID NO: 98)

Start ATG: 1

MKNRFFFSFTGYFPLGSEHCLTWICLDVLFCTASIMHLCTISVDRYLSLRYPMRFGRNKTRRRVTLKIVFVWLLSIAMSLPLSLMYSKNHASVLVNGTCQIPDPVYKLVGSI
VCFYIPLGVMLLTYCLTVRLLARQRQNLGGGQQTAAATPGWASGWLGQAPALGGPAGAAPATVGSTLSPHSAHNQGSSLTVQSDGSGYLAAPGTPCPGRRKLSISKTASVVT
WDSSRHRRRGSSFGGVRTSLLLTPTKTATTSNSSTPLRRSATLRSHQNMNYQGAGECGGKTRTTTSSPCMLQRQQTVRSHHSRNSSVISRNSSRHGRIIRLEQKATKVLGVV
FFTFVILWSPFFVLNLLPTVCAECEERISHWVFDVVTWLGYASSMVNPIFYTIFNKVFRQAFKKVLLCRYSSTSAWRPSR*
(SEQ ID NO: 99)

Name: 5-HT2 receptor-like 2
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384673
GACCTTTAAGCAGACTAAGGGGTTGAGGCAATCCTTTTTTCCCAACTCGTGTCAATTAACTCATTAAACTTCCCACCGTTTTGTGGAGCGTCGGACCCAATTAGCTTTCAAT
TATACCGTCTATGGGCTGGTTTTTCATAATCAAGCAAAATTTCAATTAATTAAAATGTAACCGAGCTAAAATGGAGGTGTGAGAAAGAGCCTTTATTACGAAGGACTCTATT
ACACATTAAAGTTTTATGGAAAGAATGAATATTCTCAGTGAGCAAGACCAGATAAGCTGCCAACATATTTTCTCTATCCTGTCAGCTTCATAAAATGTTGGTTTTGATGATG
GAAACTTTTGCCCATGAAACTTTTTGTTCGCAGATTATCGTGAAAGAGTTTTGAGCTCCGTTTCGCTGTATTGTTGGCTTGCCAAACTGTTGCATACTTTTAGGAAGTTGCA
AGGAGCGGAACATGTGTTTGCCATGTCAACTTTTTGCCATATTTGTAGCCATAACAGCTACACATCTCCGAGGAAGGTATTTTAGGAATTTTTTTTGAGTACATCAAAGCCG
ATAAAGAAAATCCCCCAGACAATGGAAGCTTACCGAAAGAAAACCGCAATGATAATTCCCCAGAGAAAAACGCCGATCGAGTACACTGAAGTAAATAGACCATTTATTTAGA

TGCTAGAATTTCTTGTTTATACAGAGAATTAATCCATTTGCACATTGAGTTACATTCAACAGCTACAGTTGGAGTTCATAGCATGGAGTTGGTTGGATTGTTAATGGGGCTC
TCAGCTGCTCACATATACGACATTTAACAGTTTAAAACGATTGACCTAACTTATGTTGCTTGTTTGCTCCGCTTGGTACTTGGTTTCGTCCTTTCATCGCACTCCAGCCTTT
AATAATTAGTACTAGCTAACTAATCGCACTCGCAGATCTGCGACACGGAAGTGGAAAGTTCAATGAAAGTTCAGTCTCTTAAGATGCACTTGCTAGGGATTTGTCTACACCG
AGTTCTCCTCGAGTCCTTGCTGGAATTCGAGAATGATTATTGGATGGTCTGTATGAGTGAGAAGAGCTGAATTCTGGTTCTCACCCTCATCAGGGGACTAATGGCCCGGGGC
ATTACGCTCAGAGTGCTGCCGGAACTCATGGCATTTGCCGCCAAAGTGGGCATCGTGTGCAGGGCATTGTTGGAGCCCCGGTGGAGTGAAAGTGGGGCATGCTGACCCACCA
GCGTGGTGGTGGTGGCCGAACTTACGCACGATTCCCGCTTGCCACGCCCATAGTAGCTGGTCAAAGGTCTGCAACGGAGTGGATGTCAAGATATCCCGTCACTTCTTAGCCA
CGATTGCCACATAACTTACCGCATACTCTCGGCTCGGGGGAACCACCGCCTTTCGAGAAGCTTTTCGTGGTATAACTGCCAAGTTTACGAAGAAATGCGCAGAAAAAGTTA
GTGTACAATAGTGAAAATATTTTTAATAGCTGCCATTCCTAGTAGCAATCTAGCATTGCCACACTTTATGATAGGTTTAACATTAAAGGCATTTGAGGAGAAATATGTACAA
CTTAGTTGTTTCCTATATAAGTTTGCAGTTTTCTCTTCTCGAGTAGTGTATTTATTCTTAAGTGTACTTACAAGCAGACTTAAAAGTTTGGACTACTTACCGTCGGTTCTGG
TTAAGCTGAATGGTGCGCGTGTCCCGCCACGTGGAAATGTGGTGTCTCAGTGCATTTCGCACCTCCGAGTTGAGGAAACAGTAGAACAGCGACACCGAGAATCCCTGTGGCA
CACACACAAACAGATCTTTAGAAAAAAAGTGGCTAAGAGGATTGACAAACGAAACCCACCTGAGTGCTGAGTAATACCGCTCGCAGGACGGCAAACATGTGTCCCATTAGGC
CAGATTCCGAGGGACCAGCCAGCACCACCAGGTAGGTGATGCCAAAAAGTGGGATTAGCACCAGCAGTGCCTTAGCAGCTTTCCTATACTGACGAGTCTCCACTGTATTGGC
CGAGCGCAGCTTTGTGATTAGAACCTAAAAGAGAAGCAACCCGATAATCTGGCTCAGTATTTTCGGCAATTTACAGTCGGAAAAATGTCGGATGATAAATTGTAAATTCATA
CAATAGCTTTAAAGTGTTTCAAGAAACTAATTGGGTGATTCATAAATGTGAGTTTTTGCATCCATGTTAACTTGTTCAATCTTCTCTGATGTTCTTTTTTTCGCATATCAAT
ATTTTTTCTGTGTAAATTCCATACGCACCCACATGATACGCAGCAGAAATGTGAGATTGATTATTAGCACCGCACAGACAGGTCCCTGGTATATCCAATCCACATGGGTCTC
CTGCATCCAGGGGCAGTTGATTTCGTACTGTCGGATGGGAAACAAAAGGAAGACTACTCTCCAGGGGTCTCTTCAAATCCCAAAATAGACATACCTTTTCGGGGGTGCTGTA
GGTGACCGTCAGACTCTTGGCCACAGCCCAGGTGACAACAAACAAGGCCGGGCCACCTGCCAGCAGGAGAAAGACTTAATTGGCTGGCACACACAATCGGGCATTCACACTT
ACCCCAGCCTATGGAGGCGTAGATATTAAAGCGTAAATTGTCCCCGGAGAAAGTCTTAACCACCAGCATGTACAGGTACAAACCCTCGACCAGCATCCAAAAGAAGTTGGTC
AGGGTGAAGAAGTGGAACAGGGTGATCAGAGCAATGCAGCTGCCCACTCCACTTCGGATGGAGATCTGTAGATTAAGTGGTAGCCATTAGTCTTAAGCTGTCAGCGTCTGTT
TATAGTTTCCTTTCACTTTGGACTAGCACCCCTTTCAGGCACGAGTATGATTTAATAAGCCAGCAATTGCAATTATCTGGCAGTCTCATAATGCTCACCTGGACGGATAACA
AGAGTATCCAGAACAAAGCCGACATGATGTACGTGAAGAACAAGTTGGCGTGGATGGTGTTGCGCAGGCAACGCAGCTCCCTGGTGAACAAAGACAAATACAAAGGACGGGG
GGTAAGTCGGGGGTTAGGCAAAAATAAATGCCATACCCAGGGTATCGGGGTCACCAGGGAGGGGGCTGGGGTGGGTACACTCACTTGAAGTAGGCGAAAACAATCAGCGCCA
GCGAGAGCGATACCAGGCTGAGGGTGTATCCAATGTAGTAGATAATGGTGGGCAGCTCCACGATGACCTCGAACTCCGGCACGGATTCTGGGGCGGGCAGGTGGGCGCAGGC
ATCGTAGTTGGTGTACTTCTCCCAAGTTCCATTGGCGTGGCAGAACCTCGTTGCATTCTCTGGTCAAATACAAAAACGCACGCAGCGAAAAATGCAAACAGTTTTTAAATAG
ACTAGTCCTATACACTTGGATCTATATTGGGAATATGCCTTCTCTGTTTGGACATCGACTTAAAAAATTCCAATTTAAATTCATAACTTAGTGCAAAGCTAACAGTTTTTTT
CAAACGCTTAGTTTAAATATAAGTCTAATGTTAAAGTATAAGGTCAAAGGGGTACTCACTGCTGCTATCGTAATGAATTCCCTGCAACTCGTCCATGCACTGGAGCACCGCC
AGGGTACCGCGAGCTGTCCTCGGCCAGCACAGGATCGAGTCGAACTGGGTGAGGCAGTGGCCACTATCGTTGCCATAGGTGGATGCCTCGATGTGCTCCTGCACCAGGCACT
GCAACTCCACGCTCTCACCGATGCCATCTAGGTTGTGCAGATCCAGGAGCGGATCACTCCCGCTGGCGTTCACCGAATCGATGTGGTTGTGGTCACTCATTCTACAACATCC
TGGCACTTGGCATCTCTGGTACTGAAATAAGTCAGTTTAAAGTTGAGGCAAAGATTGTAAAATACTTGAAGCAGTGGAAAATTTAAAATATTTTAGAATCATTATAGAGTT
TTAATACATTTCATTTAAGAACCGTCGGACAATAACGTTTTAGAAAAATGGTAAGCTTTAATCGGAAATAAAATATTATAGACTGATACGAAAATATTGCTGTCATTTCCTG
ACCATTTCCCTACGAATTGACAACGTTCATAATACTCCCCAGCTGAATTCGTTCAATCAGAAAGTTATAATTTCCTCAATTGTTAGGGGATTCCCCGGCGGTTGGCTGAAAC
TAAGCTGTCGGGCAAAGTTCAATTGCCAACCCTTCTGCTCGCTCCACCGCAGCCAGTGGCATTTCGCCTGTGTGCGTTAATTGCATTGGCCATGTTTCACCGACAACGGGCC
ACGCCCACATCGAACTCGGCTTAAAGATCCATCAACTCGGCGGGAAGTATGCTTTAAGCCAAAGACCATTTCTAGGGAAAACAGGGTGGCTAGAGATAAGTAAATGCACAGG
CCCAGTGCCATAAAAATTCATTAGTTTTGCCATCAAAGGCGGACAGAAGCGTCAGCAAATCAGAGAAGCGGCAGGCAACCTTTTATTCTGATATGTGCCCATAAAAAGCACG
TCAAAGGACCCAAGTCTGGGGTGATGGTGATAGAAATTAGAAAAGCTAACTTCCGGCCAAAGATTGTGGCCTTTGTGGCCTGGTTCTGCCTGGTTCTGCCTCAGAGACGAAA
TCTCGCATTTATGTTCTGATAATAAATGGCATAAATGATGCATCGTTACATTAATGCACGACATAAAACGGGGCGGCAAGTAAGACACACACACAAACAATACATATTAAAA
ATGTAGATGGGCCTGTTTGTACAGTGAGCTAGAAAAGTAAAAGTGCGCTCTTTAAAATTTAAGCGGAAGTGGCTTTCACTTTCTACTTAACAGTGAAATAAATTAGGCAAAC
AAGCCAGCTTGAATTCCAAGGAGAAGCCATAAGCCAAGGCAGGAATTTAAATCCACTCGGAGTGAAAATGGTTATCTCTTAAGTGATGCCCAACCCGAACCAATTACTTTCT
GTGGAATTTCCGATTTTCTCGGACGGCACATTCCCCAGGCGATTTTCTAATTGGCACAGGAATTTCTAACCACAAGCCATGAAATATTAAACCCGAAGGGAAATTGAAAATT
CTGCCCCAAGGCTTGGGTGCTCCACTTCTGGCTGCCACTGTATGCTCGGCTTCGTATTTTTCTTTCCATTGTTTCTGCTGTTTGTTTTTCTCGTAAATTAACGCAATGCATG
TACATAAAACAAAGGCTGGCTGATTATTACTGCACGGCACATGCCTGCACGTCCTTTCCGGAGTTTGCCAAGGTCCTGGGTCCTGGTTCATTAGCCGCTCGCTTACTTATAA
CACATACGCCCCGTTGCCTTCATTTTCCAGCTTCGCACGCTCACAGCAACGCAGATGAGTTAAGACAGCAATATACGCGCCGCAATTTTAAGTTGAAAAGCATATTATTTGC
AAAAAAGGGAAAGAACGGCGAAGGACCGCCTCTAATTACCTCATAAATAACGTCGCACGAAATTAGCTGGCCTAAATTATGGCCATAACTGCTAATTGCATATTTGATTTCA
ATTAGTTCAGCTGACAGGGGCCCAAAAACCCAAGAAGGACACATTTGAAATCAGCGTGCGAATGCAAATACGTTCTGACAGCTCGCTTGGACAAATCTCATTTATGGCCAGC
CAATTTGACGAATGTCCGCCCGCCGGGGAAACAGGCAAAACAAATTCAGCTTGGCCGAATCATCGCCGGCATAAAAAGGACATTACCAGAAATACATACATATACGGAGGCA
TAAGAAATGGTTGGTCCTGGACGAAGACCAACGGACAAATCGCACATCGTCCTTTCTGCTCCTCTGCCTCTTTGCAGCATGCTATTATAGAAGATTGTATGTAATCTCGTAC
GATTGCACACAATGGATTTGCCAAAGGCGACACAAAAAAAAAAAAAAAAAAAAAAATAGGGGATTGCGTGGGGGGAGTTTGCTTGTGGGTCCTTTACTTTTTTTCTCCCCCA
TGTGAGAGGTTCTGCAATTTCCTGCTGCTATAATGCTTATAAATGTTCACCCGCAATCGCCGGGGATTTCTTTAGGGAATTGAAAACATAATAACAGTGCGAATAGTGGCAA
GCATTCGCATTTTCCACTTATCCAACTAAATATGCAGACTAAAAATTAAATTCCAGTTTTTCAAGCCCTCGAAGGGGATTTTTATTTTACGCCCTTATATGTGTACCTATG
CAAAAATTCCGCAATTCCGCGGACACCGGGGAAAAGTGCTATTTTCGGTTCTCATTAATGTAGCTCATTAAATTTTATATGCATAAGCTGCCAGCTGCAAACAAAAAATGT
AAGCATTTAAAATTTATAAGCGTGCCACACAAGAGACCAACAGCACATTACGTATACGCAGCATGGACAAAGCCAGCTTGGAAAGCCAGCACGCAGGCATAAGCGTATAGAA
AGTTTGTAATTAAATTGCTCATTTAAGCGTGGAAACAATGAATTCGATCGGGAATACAGGGAGTGGGGAGTGTTCTGAGGAAATATGTAATATTCAGCGTGTAATCCGAAGA
AAATATATTACTAATTTTATTCATCCGCCTTGGGCGCTTAAATTTTATTGTTCGGACTTTTGTGGACTCTCCTCGGAGGCCGATAAGGAAAAGCCAGTCTAGACAAGCCAAC
ATCATAATTTGCAGGCACGCACTGCCATTTCCATTTCCGCCCAAGCTTGAGCCTGATCTGCTGGCCACTTTCCATGCTTAAGTCAGCACATATCCTTTAGCAGGAAGATGGA
ATGGGATGGGATGGGGCTGACACTTCAGCGATTGCCAACAAGTGTTGCGTTTTATTGTCTGGCCTTTAACAGCTCTCCTTTGCCATCCGCCCATTGTGTGTGCGGCACGTGC
AGGCGCCTGCATTTGGAGCTACGTGAGGGATCTCCACTCCTGTGGATAGTCCGAGCATAAAACCATAAAAATCCGCCGCAAAAAAAGAAACGAACCTGGCCACCAAATCTGT
TATTTATTGGCCCAATAAAAATTTGGTTTAATGCATTACAAGTTGCCGAAATGCCACGGAGATTGCCGGAGATTTACTTGGCTTGCAAAGGGAAGTCCTTAAGGCCATGGCA
ACAAGTCCTTCCTGGTTAGCAGTTAGTTTATTAGAGTGGTCAAGTGCATGGAGTGGGGGGCATTTCGATTTCGGGCTGAAATTTTCCCAATCGATACTGCAAACACGCAAA
ACAATGCCTCTGCTGCCAACTAATATCCATACTCTCGGCTCATTCAGCCACAGAGTTTTCGTTTTAGTAACCGCTGGCATTGTTTTCAGCAGGTGAGCGTGTGTATGCACTG
GCATTTTAATTAACCACAATTTGTTTCTGCATTCCAAACGCCCTGAAAACGCTTACCTTCTACTCCTAGCCAAGTTTCTATCTGCAGAATGAGATTCCTTTCGCCACTCACA
CTCACAATTTGCGTTGGAATTCTTGGGATTACTTGCGTTTAGCACTCGGACTACGGACCGCAAAAGCGCCAGGAGGCGGGTCCTTTGTCCTTTCGACCACCTACTTGACCCG
TCCTAAATGCGTTTAACGCCTTCCAAGCTGCGAAATGATTATGTTAAAGCGGTTTTTGCTGGAGCGGTTATTTTTGCTAGTTGCCTTGGTTATTGACACACGCCCATTTACG
TGGGCAAAACTTATTGAGTTCACATTTATATGCGGGATTATTTGTTATAGCGACCCTTGAACTTATTCGAAGCACTCAAAGCCAACGACTAGAAAAGTTATTCGACATTAA
TGCGCGGCACTTATTTGTTATTTCTTTTGCGTTTGCTAATTAAAATTTTCATTTTTCTCTTCGCACATTATTTAAAAATTAGTTTCCACTTTAACGGCTCTGAGGAGGACAA
TAGGTTTTCGCAAGTTTTGAAATTCCGAAACGAACCGCACTCGTACACATCCCGCTGATAACTGCGGGGATACCCGAAGGATGTGTCTTAACAAGGCTCGCCTTCTTACCAA
AAGGCGCCAATTCTGGGTTCCACAAAGTGAAAAGCTCGCAAAGCTGCATTCCACAAGCTGTTGAAAAAGCTTTCCGCCGGGCAACAGGTGGAACATCCTGACGGGAGGTTGG
GTAAAAGGACCCCAAACAGAAACCGATTCGATTGCACTTCTAATGCCTGAAAATCTGCCTATTTGATTTATGCAAATATGTGCCGGAGTCTCATGTTTGCCCAGCAATTCGC
TGGCTGTTTGTTTAGGGCATGCGGCAAAAGGGATACCCCGGACTCAACAGGATTCGGCAGGATTCGCCAGGATATTGTTTGTGAAACTCCTTGCAGGCCGTCTTTGAAGTCG
CAACCTGATTTAATTGTGATTTTCGGAGCCCGTTCAAAACACCTCTCACGTGGGCCTAATCAAATTAAACTGAAGTGGCAGCGCCTCGGGTCTGACAGTTATTGAAATATCA
ATATGGCTAGGCTCCGGGAAAATGGTTTCCCGGAAAATGTGCGAAATTAAATTGTAGAGAAAACAAGGGCAACATTTTTACTCACTGGCATCTTTTCTCATGCCTGCCAAAT
TTTATATCACAATGATAAACATTTAGACGGATGTCCTGGCTGCTTGGAATGCGTGCGACACATGAGGGTTGCCTGTGTGGATTTTTGGGCGCAAGGATATGGGCGGGGTGGT
GGGGGCGATATCTCCAACCAAACTGTTTGATTTTTAAACAAATGCTAATGTCAACTGCCACTCAGAGGCATTTAAATTCCAATTTAAAAATGCTTTCTCTCGATGCCAAAAG
CAAAGACATGTTCGTTGTCCTGTTTCGAGTCCATCCACTCGTTCTTTTTCTCTCAGATTCTCCGATTCTCTGACCCGTTTCTCCAGTTTTCTGCTCGCATTAAAGCGTTCGT
CCTGTCTCATGCATATTTAATGTCACACG

(SEQ ID NO: 100)

FIGURE SHEET 47

Exon: 7877..7337
Exon: 3829..3532
Exon: 3307..3109
Exon: 2992..2899
Exon: 2753..2579
Exon: 2520..2447
Exon: 2379..2269
Exon: 2040..1852
Exon: 1784..1669
Exon: 1420..1364
Exon: 1300..1093
Exon: 1027..1001
Start ATG: 3796 (Reverse strand: CAT)

Transcript No. : CT24513
CGGTTCGTTTCGGAATTTCAAAACTTGCGAAAACCTATTGTCCTCCTCAGAGCCGTTAAAGTGGAAACTAATTTTTAAATAATGTGCGAAGAGAAAAATGAAAATTTTAATT
AGCAAACGCAAAAGAAATAACAAATAAGTGCCGCGCATTAATGTCGAATAACTTTTCTAGTCGTTGGCTTTGAGTGCTTCGAATAAGTTCAAGGGTCGCTATAACAAATAAT
CCCGCATATAAATGTGAACTCAATAAAGTTTTGCCCACGTAAATGGGCGTGTGTCAATAACCAAGGCAACTAGCAAAAATAACCGCTCCAGCAAAAACCGCTTTAACATAAT
CATTTCGCAGCTTGGAAGGCGTTAAACGCATTTAGGACGGGTCAAGTAGGTGGTCGAAAGGACAAAGGACCCGCCTCCTGGCGCTTTTGCGGTCCGTAGTCCGAGTGCTAAA
CGCAAGTAATCCCAAGAATTCCAACGCAAATTGTGAGTGTGAGTGGCGAAAGGAATCTCATTCTGCAGATAGAAACTTGGCTAGGAGTAGAAGTACCAGAGATGCCAAGTGC
CAGGATGTTGTAGAATGAGTGACCACAACCACATCGATTCGGTGAACGCCAGCGGGAGTGATCCGCTCCTGGATCTGCACAACCTAGATGGCATCGGTGAGAGCGTGGAGTT
GCAGTGCCTGGTGCAGGAGCACATCGAGGCATCCACCTATGGCAACGATAGTGGCCACTGCCTCACCCAGTTCGACTCGATCCTGTGCTGGCCGAGGACAGCTCGCGGTACC
CTGGCGGTGCTCCAGTGCATGGACGAGTTGCAGGGAATTCATTACGATAGCAGCAAGAATGCAACGAGGTTCTGCCACGCCAATGGAACTTGGGAGAAGTACACCAACTACG
ATGCCTGCGCCCACCTGCCCGCCCCAGAATCCGTGCCGGAGTTCGAGGTCATCGTGGAGCTGCCCACCATTATCTACTACATTGGATACACCCTCAGCCTGGTATCGCTCTC
GCTGGCGCTGATTGTTTTCGCCTACTTCAAGGAGCTGCGTTGCCTGCGCAACACCATCCACGCCAACTTGTTCTTCACGTACATCATGTCGGCTTTGTTCTGGATACTCTTG
TTATCCGTCCAGATCTCCATCCGAAGTGGAGTGGGCAGCTGCATTGCTCTGATCACCCTGTTCCACTTCTTCACCCTGACCAACTTCTTTTGGATGCTGGTCGAGGGTTTGT
ACCTGTACATGCTGGTGGTTAAGACTTTCTCCGGGGACAATTTACGCTTTAATATCTACGCCTCCATAGGCTGGGGTGGCCCGGCCTTGTTTGTTGTCACCTGGGCTGTGGC
CAAGAGTCTGACGGTCACCTACAGCACCCCCGAAAAGTACGAAATCAACTGCCCCTGGATGCAGGAGACCCATGTGGATTGGATATACCAGGGACCTGTCTGTGCGGTGCTA
ATAATCAATCTCACATTTCTGCTGCGTATCATGTGGGTTCTAATCACAAAGCTGCGCTCGGCCAATACAGTGGAGACTCGTCAGTATAGGAAAGCTGCTAAGGCACTGCTGG
TGCTAATCCCACTTTTTGGCATCACCTACCTGGTGGTGCTGGCTGGTCCCTCGGAATCTGGCCTAATGGGACACATGTTTGCCGTCCTGCGAGCGGTATTACTCAGCACTCA
GGGATTCTCGGTGTCGCTGTTCTACTGTTTCCTCAACTCGGAGGTGCGAAATGCACTGAGACACCACATTTCCACGTGGCGGGACACGCGCACCATTCAGCTTAACCAGAAC
CGACGTTATACCACGAAAAGCTTCTCGAAAGGCGGTGGTTCCCCCCGAGCCGAGAGTATGCGACCTTTGACCAGCTACTATGGGCGTGGCAAGCGGGAATCGTGCGTAAGTT
CGGCCACCACCACCACGCTGGTGGGTCAGCATGCCCCACTTTCACTCCACCGGGGCTCCAACAATGCCCTGCACACGATGCCCACTTTGGCGGCAAATGCCATGAGTTCCGG
CAGCACTCTGAGCGTAATGCCCCGGGCCATTAGTCCCCTGATGAGGCAAGGACTCGAGGAGAACTCGGTGTAG
(SEQ ID NO: 101)

Start ATG: 575 (Reverse strand: CAT)

MSDHNHIDSVNASGSDPLLDLHNLDGIGESVELQCLVQEHIEASTYGNDSGHCLTQFDSILCWPRTARGTLAVLQCMDELQGIHYDSSKNATRFCHANGTWEKYTNYDACAH
LPAPESVPEFEVIVELPTIIYYIGYTLSLVSLSLALIVFAYFKELRCLRNTIHANLFFTYIMSALFWILLLSVQISIRSGVGSCIALITLFHFFTLTNFFWMLVEGLYLYML
VVKTFSGDNLRFNIYASIGWGGPALFVVTWAVAKSLTVTYSTPEKYEINCPWMQETHVDWIYQGPVCAVLIINLTFLLRIMWVLITKLRSANTVETRQYRKAAKALLVLIPL
FGITYLVVLAGPSESGLMGHMFAVLRAVLLSTQGFSVSLFYCFLNSEVRNALRHHISTWRDTRTIQLNQNRRYTTKSFSKGGGSPRAESMRPLTSYYGRGKRESCVSSATTT
TLVGQHAPLSLHRGSNNALHTMPTLAANAMSSGSTLSVMPRAISPLMRQGLEENSV*
(SEQ ID NO: 102)

Name: Diuretic hormone receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013385198
CATCGATTAACCACTATCCCATTTATTGCCTGCCCGTCTACACAAATCGGCACGTGAACGAGGACTGGTATCGACTTTATAGGACCTACGACACCGAGGGATTTGTTTTCGG
GCAGTTTTACGAGCGTCTCCAGCGGTACGAGTTGGACTGAATTAATGCCCTTTTTCAGAAAGAATAAGATTTAGACAAAATCTGTGATATCGATACTGTGGTAGCGTTGCAG
AGATCCCAATCGCGTGTGTGGCGAAGGACGCAGACAGATAACACATTCCTCCGTGCTGAAAATCAATTATAATTAGTCGTGGGAAACAACTAAAAGAATATAAAAATGGCCC
GTTGGTTTATAAAAAATATAATACTTTAATCTGATCTCTTATAATGTAAACTATCATGCAATACTTAGGGACGAACGGATGTGGTCTACATGTATGTCGTATACGTATATTT
TTGTGTGGTGTTCATTCATAGTGTCCAGTGGGTATAGGACTTGTTGATGTCAGTCTACAATAAACGATAAACAGAAAAAGGAAAACCCCAATCAGAAACATAAGAAAATAAG
TAAGAATAAAAAGGATATCTAAAGCCAACCAACCATGCAGAATAAACATAACATTAAGAACTGATGTAGGGAAATGCACGGTTTTAAGTAGATCACTTTAGTTGGCACGAGT
GTGTTAGTATATTTAGAGAACGGGTGCAATTTAACTTTGGAATTATTTGGATGGGTGGATGTCAAAGAAGTAGAAAGCATATACATGAGCATTGCGTATACGCACCGTAGAC
TTTCGGTTCGCGATCTCAGCGAGTAGTCTTTTGAGGCGCGATGCCTGAAATGGAAATTAATTGGTTAAATGACAACAACAACAACAACATCGACGACAATAGCGAACAA
CTACAATATCAGCCATAGCCACATCAACTAGGTAGTGCAAATAGTGAAAAGCCAATATGTCGACAGCCAGGGTTATATAAATATAATCCAACAACAATAATGTGCTAGAATG
CCAAATGAGTGCAGTCGCCCCTCAATGATTGGTTAGCCCACCTGCGATTCTTGATGGAACTGTTCCGATGGATATTCCTGCTCTCCCGCCACCGGGTGAATCCATGCCTCAG
CGTCTGGCGCACCTCTGAGTTAAGGAAACAGTAGAACAGAGCCACAAAGAAGCCCTAAAGAATGGATTTTCGAAGATCAATGTTTATGGAAAGACTGTCACCTAATCTATAT
ACCTGCGTGCTTATGAGAAAGGCTCTTATGGCCTCGAAGAGATTACGACTGATACCCTGTTCCGGGCCTGTTAGCACCAACAGATAGGTGATGCCAAAGAGAGGTATCAGCA
CCAGCAGCGCCTTCGAGGCCTTGTAATACTGCCGCGTTTCCAGGGTATGAGCAGAACGTAATTTAGTGATGAGTACCTGTACGATATGATATAATAGCATGAAGCTAACTCA
CTTAAGCCAGAAGTTAACCGGGGGACTCACCCACATGATGCGTATGAGGAATACTAGATTAACCAGCAAAGCCAGTGATGCAGGTACCTTGAATATCCAGTCAATGTGAGAT
TCACGCATCCATGCACAATCAATTTCCAGCTTTTGAAAAATCAGAAGAGGTTAGCAAGCATATCTAGGTAGGTGGCAATTGGTATCATACCCCATTGAAGTGCTCGTTCTCG
AGATGCGGAGCAAATGCCTTGGCAATGGACCACACCAAAATGCATACGGCTGGACAGCCCCAGCCGATGAGGGCGTAGATAATAAAGCTAATGTTATCACTGGAGAATGTTT
GCACCACCAGCGTGTACAGATAGAGGCCTATAAAAAGCAAATGATAAAACCGAGGAAATTCAATTGATACCGCCTTTGAACCCACCCTCCACAAACATCCAGAAAAAGTTGG
TTAGGTAAAAGTACTGAAACATGATTACCAACGTTATGCAGCCAGCCTGACTAGACTCTGTGGTTATCTAAGTGAATAATTAATTAGTGGAGTGGGATATATATATTCCAGC
TTATCTTACCACTTGCAGGAACAGTGTGAGTATCCAGAGGAGTGCGGATGTGATGTAGGTGAGGAACAAATTGGCATGAATGGTGTTTCGCAGGCAACGAAGATCTCTGAAG
AGTTTTAAGCATGAAAATGATGCAACTCATGTGTGGCATTACAACTCACTTAAAGCTGAGGAATATGATGAGAGCCACCACCAAGGTGGCGAAGCTCAGGAAATAACCGCCG
GCATATATGATGGCCGGCAGTTCGACGTTGGGTGAGAAGTCGGGCACCACCGGTATGGAGCCCGAGTTCTGGTGACAGCGGTCATAGTCCGAATAGTGATCCCACGTTCCGT
TTGGAAAGCAAAAGCGGGTGGCATTGTCTGCAAATATATTGATTTTATTTAAGATCTAAACCCAATCTGGAATACCCACCTGTGGTGTCGTAGTGCACGCCCTTGAATTCCT
CGAAACAGGGTAGTACAGCCAGACTGCCGGCGTTTGTGCGTGGCCAGCACAGGACCGAATCGAACGAGCTGGGACATTGTAAAATCTGTAAGAAATAACAATTGAGAATATA
CAATGATTGGCAAGGTAAACTTCGTTGAGCATTTACTGCAGTATAATTCGTGTTTCCCCTCTGTAAACGCAACTGACCTAGCCCAAAACCCATTAAATCAGAGTACATAAGA
TTCATGCCCCGTTTCTGACTCGGTATAATAATGATGCCTTCATTGCGGATCCTGGTCATTTATAGCTCTGACGCATGATTTTCTTTAATCAAATTTAGTTTCCAGACTTTAC
ACACACACATACGCAGCGTGAGCCAAATATCGATGTAGTCCCAAAAAAAAAATAAAGAAAGAATCACCAGCTACCGCTTTATTTATATATTTGCAATGGCGTGCTAAACAAT

TTCGTGCATACGGATACACAAGTTATATGTACATATATGTATATTAACGTTAAGTATACGACGACCCGGCGATTGTTTACCCAGTTCAGTTCGTAGCGCCAAATTAATCGCG
ACGCGCTATGGTAGCGGTGTTAGAAACAGAGCTTCATTGATATTTTTTTATTGACCTTATCAGGTGACGCCGAGAGCGCGGAGCGGAGACTCGTCGAAGAATTCACCAAAT
TGGCAATTCTTTTATGGCCGGTCTTCCCATTACTGATAAGCCCGAGAACCTTTTTGCGACAGCTAATGCGATCGGACTTAGCCACGCGACAAGTGAGCTGTCAACGAGTAAG
CCAAACACTGGAGACTGGAAACTCCGGCAATATTTGTTTCTTTGCTCTTCTGATTTTCACTCATTAGCCTCGTCATTTGGACACAGTCATAGTACGAATACATTTTATGACA
TTTTCGAGGAACAACGCCCTTGATCGTGCATTCGGTATAAACAAAAATGGCTAAATGCAGCCTCAAGTGGATTTCCTTGGGTCTAGTTTGTTTACAGGCGAGAGGAATACA
TAAATAAATATGTGCTGCATTTTCCGGAAATAAAATGCCAAAATAATTTAATCCGAAAGCGATCGACAGTTTACTTGCCTGGTTTTTCATCGTATCGTACTTAAAGTGAAAG
AGATTGACGGGCGGATGGATTGAAGCCAATTATTTCGGAACAGAAGATGGATCAATTAGCGGGTGCGATTAATTGGGAAATTCCTCAATCGGCAAAAGAGATGAATTCATTG
GCCAAACAAAGGGTAATTAGGGAATATTTATTGGC
(SEQ ID NO: 103)

Exon: 2731..2708
Exon: 2549..2432
Exon: 2379..2178
Exon: 2122..2026
Exon: 1971..1878
Exon: 1819..1659
Exon: 1597..1487
Exon: 1420..1235
Exon: 1174..1001
Start ATG: 2731 (Reverse strand: CAT)

Transcript No. : CT24959
ATGAAGGCATCATTATTATACCGAATTTTACAATGTCCCAGCTCGTTCGATTCGGTCCTGTGCTGGCCACGCACAAACGCCGGCAGTCTGGCTGTACTACCCTGTTTCGAGG
AATTCAAGGGCGTGCACTACGACACCACAGACAATGCCACCCGCTTTTGCTTTCCAAACGGAACGTGGGATCACTATTCGGACTATGACCGCTGTCACCAGAACTCGGGCTC
CATACCGGTGGTGCCCGACTTCTCACCCAACGTCGAACTGCCGGCCATCATATATGCCGGCGGTTATTTCCTGAGCTTCGCCACCTTGGTGGTGGCTCTCATCATATTCCTC
AGCTTTAAAGATCTTCGTTGCCTGCGAAACACCATTCATGCCAATTTGTTCCTCACCTACATCACATCCGCACTCCTCTGGATACTCACACTGTTCCTGCAAGTGATAACCA
CAGAGTCTAGTCAGGCTGGCTGCATAACGTTGGTAATCATGTTTCAGTACTTTTACCTAACCAACTTTTTCTGGATGTTTGTGGAGGGCCTCTATCTGTACACGCTGGTGGT
GCAAACATTCTCCAGTGATAACATTAGCTTTATTATCTACGCCCTCATCGGCTGGGGCTGTCCAGCCGTATGCATTTTGGTGTGGTCCATTGCCAAGGCATTTGCTCCGCAT
CTCGAGAACGAGCACTTCAATGGGCTGGAAATTGATTGTGCATGGATGCGTGAATCTCACATTGACTGGATATTCAAGGTACCTGCATCACTGGCTTTGCTGGTTAATCTAG
TATTCCTCATACGCATCATGTGGGTACTCATCACTAAATTACGTTCTGCTCATACCCTGGAAACGCGGCAGTATTACAAGGCCTCGAAGGCGCTGCTGGTGCTGATACCTCT
CTTTGGCATCACCTATCTGTTGGTGCTAACAGGCCCGGAACAGGGTATCAGTCGTAATCTCTTCGAGGCCATAAGAGCCTTTCTCATAAGCACGCAGGGCTTCTTTGTGGCT
CTGTTCTACTGTTTCCTTAACTCAGAGGTGCGCCAGACGCTGAGGCATGGATTCACCCGGTGGCGGGAGAGCAGGAATATCCATCGGAACAGTTCCATCAAGAATCGCAGGT
GGGCTAACCAATCATTGAGGGGCGACTGCACTCATTTGGCATTCTAG
(SEQ ID NO: 104)

Start ATG: 1 (Reverse strand: CAT)

MKASLLYRILQCPSSFDSVLCWPRTNAGSLAVLPCFEEFKGVHYDTTDNATRFCFPNGTWDHYSDYDRCHQNSGSIPVVPDFSPNVELPAIIYAGGYFLSFATLVVALIIFL
SFKDLRCLRNTIHANLFLTYITSALLWILTLFLQVITTESSQAGCITLVIMFQYFYLTNFFWMFVEGLYLYTLVVQTFSSDNISFIIYALIGWGCPAVCILVWSIAKAFAPH
LENEHFNGLEIDCAWMRESHIDWIFKVPASLALLVNLVFLIRIMWVLITKLRSAHTLETRQYYKASKALLVLIPLFGITYLLVLTGPEQGISRNLFEAIRAFLISTQGFFVA
LFYCFLNSEVRQTLRHGFTRWRESRNIHRNSSIKNRRWANQSLRGDCTHLAF*
(SEQ ID NO: 105)

Name: Diuretic hormone receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013385192
TTTCCGCTGCCATCTGCAGCTGTAGCTGCCTAACAAATGCAATTAAAATCTTTGCGCAAGAAAATTGTTGCAGCAGGCACGCAGAGATACTCGTACTGTATGTGCTGCAATA
GTTTGGTGGCAGCAACGGGGCGGCAGATAACATATTGCCCACGTCCGGCTGAAATTGTACGGACAAGGCGAACTCAAAAAAAAAAAAAAAAAAAAAGGAATCCCTCAAAAGGA
GAACCAAGAAAAATGGGATTTATCTCTGGCCAAGCTCCAGCCATACAAATCTAATTTAGACCACACGACGGCTTCAGATATTTTATATTCCCCACTCAGTTGCGGTCCATCC
TGATCTCCTCCAGCTTGCTTGATGAGCCCAGCCGTTGGCCAAATGAAATTATTTATGGGCAGTCCTAGGTCAGTGAGAATTTGGTTTTTTAAGGGTTTGAACTTTGCGGCTC
AACAATGCCTGTGATTTTTTTTTTTTTTATTATTATTTTTTGCATTGCATCCATTAGAGAAAGCACTCCTTTGGCATTCTACTCAGCTACTTGCCACATAAATATCTACGGAGG
GTAGCTAACTAGTCCTTATCTCAACTTCCCCATTTATCCTCGCCGCAAGTGCACTTTTCCTGCTGCTAATTAAGTTGCAAGGCAGCTTTCTCAGCAGGCGTTGCACGAAGTA
GATAAAAGGAAAAGCCCCCGATTTGCAGACACTATGCTGGGAAAGATATATATATATATATATATTTTGCCAGCATTCGCTAATCGGCTTAGGGCAGCCACCATGCGACCACAA
CAACAACAACAACAGCAGCAGCAGTGTGCTCCATGCAATATTTAATCACATTAGTTGATGGGGTAACGGCCATGTACGCAACAATTTACAACTTGCATTAATTATACGTCTA
GTTGGACAATGGCGCCTCTAACTATCCCACTCTAAATGTATCTCTATCTCTATCCTTATTTGCAGGGGACACCCGTTGATAATAAATTATTAATGCATCTGCTGGGAAGGCG
GCACTGCACGGCTGTCCTTGGCAGCAAATCCAGTTGCAACATCGGTGAGTGACGGGCGGCAATCACAAAGGACTTGCAACAGTTGCCGTTGCTTCTGCTATTGCTGTTGCTG
TTGCTGTTGCAACAATGGCCAGTGGCAACAATGAAACTGAGCCGCTCTACTGCGGCAGCGGCATGGATAATTTTCATACAAGGTAAGCGACTTTCACCAATGTCACGATCAA
GGCCAAGGGGCTTTTCCACCAGACATGGCGACAACAATACCATAAGCCAACAGTTGGGGGCGTGGTCGGGGGCGGAGGCGATGGTGGCAACATCATAGCCATTAGCTGCTGG
CGCAAGGGAACCGTTCAAAAATCGATTATCGCCCCATTTCGGGGGAGCTTCTATTTTGATTTGCCGTACAATTTTCTCGGGCGATTAAACGACGAAGCAGAACGAAAACAAA
AAACAGATTTGTCAACAGCAAGGTCAACAATTGATGGCTGAAATCAATTTAATTGACCATATCCTACGGGCCCTCCAAGTGGCCATCTGCTGCACCTATAAAAAAGTGAATC
CGGTCTGCGATTATTTATATATTCGTTGCATGGCAGGCGGTCGTAAAACCTCGAGATGATGATTAAAAGCGGCCCTAAAAACTTAATGGCGGTTTAGGAAATTCAATTCCTG
TAATTTAAGCCGAGTCACCATTCTTCGAAGTTCTTACATGTAAGCGATAATAAATAGTTAAGTCAATTGGCCAATAAACCTATTAATATTGTGCATTTACCACGATTAGACT
TTGATTAAAGTGACAATGCTGATTTCTGTAGAGGAAATCTAGTTCTAGTCTTCCCACAAAGCTATTTAGTTACTCTTGAATAAATATGTTACTTTTCTTTTGCCAAAACCAA
CAGAATTTTAAATTTAATAATTTGGATTTTTTGCAATAAACTGTACTGATTAATGGGCCACACAAAAATGTCTAGTTTATTATGGAGCTCTTGGTTTCATAAATTAAGAACA
TAATCCAATCGGCATATAAATCATTGATAGCAATTTATTTTCCGTGATGAAACTGTGCTCCGTGTGAACTGCGAATTACTCATTCTACGGTTGCAAAAAAAGCCACCAACGG
TCAACATTTAGACCAGGACTTTTAGTTTTAATTAGAGCCAGCCTGGCCAACAGCAGTGTTAATGACCACAAAGTGGCTGGCCACAGGATCAGCATCCCAGAATGCGATGCCG
CATTTGCTTTAATTAAAGGTAGTAGCTGGAGTTTGAAAGATGACTGTATGGCAATTAGATGTGTAGCCAGAACACTTGGCCATTTACTTTTGTGTCAAAGTCGTGCCAAATT
GCCAGCGGAGGCGACACTTGACGCTGTCACGCCCCAGACAGACGCAGACCGGCCCAAAAGCACCCACTCAGCCGTCTCCAGGCGCCACTCAAGCGGCAAAGGAACGCCAAAA
CACTAGGACACAGAACGCCAGAAGACTCGAAAAAAAAGTATAAAATTCAAACGCGTATACAATTCAATGTGTCACATCGAAGAAGAACCCTTTCTCTGGCGGCAACGTGACG
TATACGCAATATTAATGGCAAAGCGAAAAGGGCCTCAAAAGCTAACGAAAAACCGCAGCATAACACGCACAAAACAAGCAATTAACATAAATAATAAAGAACGCACAAAGGC
GTTCGGCTCTCCGTTGCCAAAGAGAAAAGCCAACGAAAGGAATAAAAAACGCCAAAAAGCCCGAAAAAAAACCGAAGACCCAAAAAATGTGTAAAATTGATTTCGGAAAAT
ATTTAATTTACCAAATGCATGCTCGCTTTTTGCGGCCGTGTACGCCCCGAACAAATAAAACAAATAGGCACACATCGTCTGGCATTTCATGTTAAAGGCTTCTCATGCTGAT
GGACAGGATACAGAAAATATATATACGAAGTGAGGTGTGTGTCTACGTGCCACAAAAACAGCGAAATTGAAACGCTTTCCGAATGCCGGCGACACTTTTTGGCGGCACAACC
CTCGTCTTGTCGACCAACTCCTGGGGAATTTACAACTCCATTCATCAAACTCACAAAAACTTTGCCTGGTGTACGGTCTTGGGTTCGTCGTCGCAGGATTTCACCTGGCTAA

AAATCCAGACTCCCAGCCATCCCTGAGTCTAATTTATGGCCTTTGATAATCCATGTTGATGAATGGCGGCAACGACAACAACGTCGACGACATGAATGAAGTCCTGGAATTG
TTTTGCACCAGGATGGCATCGGGGCTCCAGGTGGGACGTACTGGCTCAAAGTTATTGGCCCAGAAATCAGGCATAGTTAGCTGCCGAAATGAAACCCAAATACCGAGAAAAC
TAGGCAAAACAAACAGTAGTACACCGGAAATGCATATCATTGTAAAAACTACATCAGTTTACCTAAAAGGCTTGGCTTTTAAGCTTTCACATTTATAAAATATTGAAAATGC
ATATAAAAGTATGAAATTAATTCCCTTTTGTCAATAAACTTTCTTTCTTTCTTTCTGTGTAATATGGGGGATACCGGTTTTTTTTTTTTTTCAATGAAATCCCTTCGAAAGG
TATAAGTTCAGAATCGAGAGTTTTATGCCAAGTTGGGCACAGTTTTTTTTTTCCCCAGCTACCTAAAATAATAGAGACATTTTCCTCCCACTACAACTGATTGCATTGCCGG
TGCAGAAAGTTTTTTCAGTTGGTTCGGAAAAATTTGGTTCGCAAACAAATTAATATGAACTGGCAAGCATTTTTCGGGCAAAAAGCTCTCATCTATGTAGATTGGAATGGAA
ATTCCGGCTAGAATTGCATAAGACCACCTGCAGTGTGGGCTAACATGACTAAAAAGTTGTCCACAAATTTGGCTTAGATTCTCCAATAAAACTGTCGTTCGGCCAGGAATCC
CCTTTTTTGTTTCGAGTGAATGGGGAATTTCGCACGACAGACAGCAATAAAGAATTTAACTAAAGTCCTGACACCGACAGCACCAGCAGGACGCACACGTGTCACTCCATTT
GGAGAGCTTGGAGTATATTAAACATTTTTCCCCACCAGTCAGCCGCAGGACTTGCATCGGTCTCGCCTCGCATTTTCCTATATAAATTTTATGCTAAGTCTAATTTGTTGG
CTGCAACTTGCACAAAGGCAAAAAATAAACAAGGGCGAAATGCCGAAAGCCAAAACCCAACCGAAACCGTTGAGGGCTGCCTCGCTTTTTTCCTGTGCCGAATTCCCTAAAA
CTTTGCACATAAATTTGAGTCCTGCGCCTGGGCTTTTCCTCTTCCACCTTTTTTTTAGCCTCAAAGCGCTCGACGAAAACTAAACAAAAGCTGGTAAAATGTTTATCCCAG
AAGGGGGCGGTGTGGGCGTGGCTTGCTTAAGCCCTAGACACCCGAGAGTTTATTGCCTCAGTGTTGCAAACGCTGCTGCTTGGCAAGTCAACGGAAATGGCTTTCGGATGGT
TTTATTGCCTACTTTTGCTGCGAGAAAGCGGACTAGACATGGAAGTAACAAATTGGTTAAGGCTAACTAGAGCATAATAACAGAAACATCAGAAACATAATTTGGGAAATGC
CTGCCACTCACCGACAACATTTCGATTTGTGTCTTTATTACACTTCTCTGCCGTGATGGGTTATAACTTTTGGATCCTTGCAGAGAGTTGGGAAATTGCTTAAAAATAAAAT
TTGTTTAATTTAATTTGAGGTGGTTTTTCAAACAATTTTAATGGATTAATAAGTTTAGGAAGTTAAGTATATACATCTATCTATTATACCATCTTCAAATATTGTTGAATTG
AAGAGATAACTCGGAGCTAATCCCATATTTTTCATTTTTTTCAGCTACAAGAACATGCATGGCTATGTTTCGCTGGTGGTCTGCATCCTGGGCACCATCGCGAATACCTTGA
ATATCATTGTGCTAACCCGACGGGAGATGCGCTCCCCCACGAATGCCATACTCACGGGTCTGGCCGTGGCCGACCTGGCAGTTATGCTGGAGTATATACCCTACACCATACA
CGACTACATCCTGACGGACAGTTTGCCGCGGGAGGAGAAGCTCAGCTACAGCTGGGCCTGCTTCATCAAGTTCCATTCGATTTTCGCCCAGGTTCTGCACACCATTTCCATT
TGGCTGACGGTGACCCTGGCTGTTTGGCGTTATATAGCGGTGGGTTATCCGCAAAAGAATCGCGTATGGTGCGGTATGAGAACCACCATAATAACGATAACCACCGCTTATG
TGGTGTGTGTTCTGGTGGTGTCGCCGTCGCTCTATTTGATCACGGCTATAACCGAATATGTCGATCAGTTGGATATGAATGGCAAAGTGATAAACTCCATTCCCATGACCCA
GTACGTAATCGATTATCGTAATGAGTTACTGAGTGCCAGGACGGCTGCCCTGAATGCCACGCCCACCAGTGCACCACTGAACGAAACTGTGTGGTTAAATGCGAGCACCTTG
CTGACATCGACAACCACCGCTGCACCACCCACGCCATCGCCAGTGGTGCGAAATGTTACTGTCTATAGGCTATACCACAGCGATTTGGCGTTGCACAATGCCTCGCTGCAAA
ATGCCACATTTCTCATATACAGTGTAGTGATTAAGCTGATACCATGCATAGCACTCACCATTCTGTCGGTTCGATTGATCCTGGCCTTACTGGAGGCCAAGCGGCGGCGGAA
GAAGCTCACCAGCAAGCCCGCCACTCCGGGTGCCAGTAATGGAACCAAATCACCGGCCAATGGTAAAGCAGCGGATAGGCCCCGGAAAAATAGCAAAACTCTGGAAAAGGAA
AAGCAGACGGATCGCACCACGAGAATGCTGCTGGCGGTGCTACTCCTCTTTCTCATCACTGAATTTCCACAAGGGATTATGGGTCTGCTGAATGCTGTGCTCGGAGATGTCT
TCTATCTGCAGTGCTACCTAAGACTGAGTAAGTAAATATGTACTATGCATAACGCCACAGTAAATTCCACAGATTCCCATTTCACTTTAAGATGTAATAACCCAATGCCTTT
TGCCTTTTGCCACCAGGTGACCTGATGGATATCTTGGCCCTGATCAACTCCAGCATCAACTTCATTTTGTACTGCTCCATGAGCAAGCAATTCCGCACCACGTTCACGCTGC
TCTTTCGTCCAAAATTCCTGGACAAGTGGCTGCCGGTGGCGCAGGACGAAATGGCAGCTGCTCGAGCTGAACGCTCTGCGGTGGCACCGGTCCTGGAAAAGGGACGACAGCA
GCCGCAGGTGGTGATGGCCAGCACGACCACCAACATCACGCAGGTGACAAATCTGTAGCACAGGAGGAGTCGTGGTCGCCGAACTTTGCTCAGTCGCCTGCTGAGTGTCCTG
AAACGCGGCAGAAGGCGCTCCTCCGGCGAAGGAGGAGGCGTGGGAGGAGGTGGTGCGCCGTTGGCCGGCAACGATGCGGTGGAACCAGCGTTCCAGGCCATCGTGGTGGTG
TGGACAAGGTGAGCGGTGCCACGGAGAATCAGCTGTACACCGCCGAGCAAGCTCGTATTGTGACGTAGTAACCCTGATAGTAGTACATATTATATGCCTAGTTATCCTTGTA
AAAAGTAAACTAGTTGTGTAAGAATCGCAGATAACCACAAGTTATGTGCATACTTTGGTGGGAAATCGTAAAACCCGTACAGTCATAATAATAAAATGCTGGTTCATCACTC
ATACGCACCGTTTGACTTCACTCTATTCCACATTCACAGTCAAGTAGGAGCTGTAAAAATGCTGTGCATAAACGAGCTCGTATACGCTACAGCTGGCGTCACAATAATAGTT
ACCATTGGGTTGGCTTCACAACATAGATCAAAGTCAAAAAAATATATAGTTTTGTAAAATGTTTCATAAATCTATCTTCATATTTAACTAACACAATTCAGAAACTTTCGCTT
GCATGGAAAATGTTTTATTTTCCCCACTTGACCCAATTGTGCGCAACTGTTCAGCTTTTCGGCATGAATCAAATCAAATATAGTAAACTTTTGCATCTTTTCCACTTTGATC
CCATGCGTTTTCATGTTGATTTCATTCGCATAAAGTATGCAGCACACGTACTCACTCAAAACTACGAATTTTCGGCAACGATTTGAGCAGAATATTTGCCATTTGTTGTACA
GATTATATATACACACACACACACTTATGTCCGTATCTTTAATAATGGCAATTAAGTCCTAGAAATTTTGACCTGATATTGATGCAGGCATTGGAAAAATAAGTGGCAGTTG
CAGTGCGGTGCGTTGATAAATGAAATACTTTGCGAATTCTATTATTGGAAATCGCAATTTGGGCTGTGAATGAAATTATAGTTTCTTGAACGCAATTGAGAATTCAGTTCC
GTTTCGTGGGTTACTGAATTTATGGGAAATAATTTAAGGCAGTCACTGCAGCTGACCGAATAATCCAATGAAACATTGCATAAACAGTTTGCAGTTTTTCCACACTTGCGAA
AATTGTAATTGCAAATCGCTTAAGCCCCACGACGAATGACAAATGGCCAGCGAACGCATA
(SEQ ID NO: 106)

Exon: 1001..1202
Exon: 4861..5963
Exon: 6065..6282
Exon: 6331..6564
Start ATG: 1135

Transcript No. : CT25824
GGAAGGCGGCACTGCACGGCTGTCCTTGGCAGCAAATCCAGTTGCAACATCGGTGAGTGACGGGCGGCAATCACAAAGGACTTGCAACAGTTGCCGTTGCTTCTGCTATTGC
TGTTGCTGTTGCTGTTGCAACAATGGCCAGTGGCAACAATGAAACTGAGCCGCTCTACTGCGGCAGCGGCATGGATAATTTTCATACAAGCTACAAGAACATGCATGGCTAT
GTTTCGCTGGTGGTCTGCATCCTGGGCACCATCGCGAATACCTTGAATATCATTGTGCTAACCCGACGGGAGATGCGCTCCCCCACGAATGCCATACTCACGGGTCTGGCCG
TGGCCGACCTGGCAGTTATGCTGGAGTATATACCCTACACCATACACGACTACATCCTGACGGACAGTTTGCCGCGGGAGGAGAAGCTCAGCTACAGCTGGGCCTGCTTCAT
CAAGTTCCATTCGATTTTCGCCCAGGTTCTGCACACCATTTCCATTTGGCTGACGGTGACCCTGGCTGTTTGGCGTTATATAGCGGTGGGTTATCCGCAAAAGAATCGCGTA
TGGTGCGGTATGAGAACCACCATAATAACGATAACCACCGCTTATGTGGTGTGTGTTCTGGTGGTGTCGCCGTCGCTCTATTTGATCACGGCTATAACCGAATATGTCGATC
AGTTGGATATGAATGGCAAAGTGATAAACTCCATTCCCATGACCCAGTACGTAATCGATTATCGTAATGAGTTACTGAGTGCCAGGACGGCTGCCCTGAATGCCACGCCCAC
CAGTGCACCACTGAACGAAACTGTGTGGTTAAATGCGAGCACCTTGCTGACATCGACAACCACCGCTGCACCACCCACGCCATCGCCAGTGGTGCGAAATGTTACTGTCTAT
AGGCTATACCACAGCGATTTGGCGTTGCACAATGCCTCGCTGCAAAATGCCACATTTCTCATATACAGTGTAGTGATTAAGCTGATACCATGCATAGCACTCACCATTCTGT
CGGTTCGATTGATCCTGGCCTTACTGGAGGCCAAGCGGCGGCGGAAGAAGCTCACCAGCAAGCCCGCCACTCCGGGTGCCAGTAATGGAACCAAATCACCGGCCAATGGTAA
AGCAGCGGATAGGCCCCGGAAAAATAGCAAAACTCTGGAAAAGGAAAAGCAGACGGATCGCACCACGAGAATGCTGCTGGCGGTGCTACTCCTCTTTCTCATCACTGAATTT
CCACAAGGGATTATGGGTCTGCTGAATGCTGTGCTCGGAGATGTCTTCTATCTGCAGTGCTACCTAAGACTGAGTGACCTGATGGATATCTTGGCCCTGATCAACTCCAGCA
TCAACTTCATTTTGTACTGCTCCATGAGCAAGCAATTCCGCACCACGTTCACGCTGCTCTTTCGTCCAAAATTCCTGGACAAGTGGCTGCCGGTGGCGCAGGACGAAATGGC
AGCTGCTCGAGCTGAACGCTCTGCGGTGGCACCGGTCCTGGAAAAGGGACGACAGCAGCCGCAGGTGGTGATGGCCAGCACGACCACCAACATCACGCAGGTGACAAATCTG
AGTGTCCTGAAACGCGGCAGAAGGCGCTCCTCCGGCGAAGGAGGAGGCGTGGGAGGAGGTGGTGCGCCGTTGGCCGGCAACGATGCGGTGGAACCAGCGTTCCAGGCCATCG
TGGTGGTGGTGGACAAGGTGAGCGGTGCCACGGAGAATCAGCTGTACACCGCCGAGCAAGCTCGTATTGTGACGTAG
(SEQ ID NO: 107)

Start ATG: 135

MASGNNETEPLYCGSGMDNFHTSYKNMHGYVSLVVCILGTIANTLNIIVLTRREMRSPTNAILTGLAVADLAVMLEYIPYTIHDYILTDSLPREEKLSYSWACFIKFHSIFA
QVLHTISIWLTVTLAVWRYIAVGYPQKNRVWCGMRTTIITITTAYVVCVLVVSPSLYLITAITEYVDQLDMNGKVINSIPMTQYVIDYRNELLSARTAALNATPTSAPLNET
VWLNASTLLTSTTTAAPPTPSPVVRNVTVYRLYHSDLALHNASLQNATFLIYSVVIKLIPCIALTILSVRLILALLEAKRRRKKLTSKPATPGASNGTKSPANGKAADRPRK
NSKTLEKEKQTDRTTRMLLAVLLLFLITEFPQGIMGLLNAVLGDVFYLQCYLRLSDLMDILALINSSINFILYCSMSKQFRTTFTLLFRPKFLDKWLPVAQDEMAAARAERS
AVAPVLEKGRQQPQVHRRSRGRRTLLSRLLSVLKRGRRRSSGEGGGVGGGGAPLAGNDAVEPAFQAIVVVVDKVSGATENQLYTAEQARIVT*
(SEQ ID NO: 108)

Name: G-protein coupled receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384565
TAACCTTCTTTTGACCTTCGTAAGTCTTCCACTGGCACTTGAAAAGCTTTGGGAAAACAGCTGTGAATTTGCCCAGCAGAGCTTCCGAAACGACGGCTTTTAGTACCTCCGC
CGTAACACCACCCACAGTGGGTCGGTTCTGGTCCATGGTGGGGATGGTGTGGTGGTTCCGATGATACGGGAATCGCTGGAGAAGCGTTACTGATGCAATGCACTGGATACCG
TTCCTGTTCAGATCATTGAGTACTCCAGGCGGCATTTCTATTGTGGTTACAATAGTAAGTAAGGAGTAGTAAAAGTTATTCTGAAATTATATAAACGGCTTGTTGGGGTTTT
AAAAATGTTGCCGGTTTTAGATTTAGCAGTGACTCCAACCATCTGATGGGGAACTTTGAGCAAGTGAGCTCCAATAGTTTTGGCGTCGCTTCCAAAAAGCTCTGGAAAAGCT
TTTATCAACAGGCTTGCCAAAATCCGCTATACTAGCATATGTAAAAGAAAAACTTGTATACCCTACCAATAGGCATACATCGCATCAGGACTTATTTAACCAACAAAAATAA
AACACAACATGAATGTACATTCTGGTTTAATCGCTTTATTTAACTTCAACTTTGAAATTCATGTCGCATGGCTTAACTTGGTTTTCACAGACGTTCAGGTATACATAAGCTG
ACATATACACAGTAAATACTCTAACAACATTAAACATATGTATGTACATAATTAACTAGTCCGTAGAAACGTATTATGACCATACACTCAAAATAAATGCGCATAGCTCCTA
ATCTCTATATTTCACTACCGAACTAATTGATTTAGATCCGATATAGAATTTGACAACGTTTCCGAAAGAGAGCCTAAGAGAGCTGGAAAGTACTTAACGTAAAAGCTTTGAC
TTGCGAGTATTTCCTCAGGGGTCACCCCTTACCCCTTTCGCATCCTTGTGATCCTCTTCCAGTGGACTTTCCTGCTGATCCCGCTTCTTTTCCGACCGCGAAGGTCAACTCT
CCGCGCTGGTCGCCAGAAGGGACGGCGAGGAGGTGGTCAGACCCACTGGAACCGATTCCTCCTGAGCCTGGCCCATGGGCGGTGGTGATGGAGTGGGCGTGGAGGCCGTGGA
CGACGCCGTCGATGAGGTGGGACGGGCGGCGTGAGGCTGGAGGAAGTGAAACAAGTTGGAATGCTGCATGAAACTGCCCACCGATCTCAGCGGACTAAAAAAGTTCTCTGTG
GGATGGTTGGCTGGCTGCGGATGCTGTGGAAGTACAGAGGACTCCTGATCCCTGCAGAGATTACTGCTGGGCTCGTCGTGGACCCCTCGTCCCTTAGCGATGGAGCTACAGT
TGCAGGATCGGCTCTTCTTCAGCGAACTCCGAGTGAATGCCGTGCCCTTTCGATACTTAGGCTTGTGCCCATTTGAGCCGGGACTCGCCTCACTATCGATCACAAAGATACC
GGATTGCTCCTGGGGAGCAGGACTCGCCGTGGATCCTGGCAGGCAGTGAGCGCTATTGGAGTTGCCAGCAAGCAGTTTGGCCCTCAGAAGAGCCGGAAGTGGGCCATTTGTG
GGAGATGTTTGAGTGGAGAGCGGTCTCTCCCTGGACTGTTCATCTATGGGAAGATCCTGATTGTTGCCCATTGGCCTGGAGTCCGTGTCATCCGATCGCAATTTCTCCCTCA
TGACCAAAGCCAGGGAGCGACGCAGTTCCACGGGATCTATGCCAGGCTGAACGGGATTTGGAACAGAGCTGGCCGGCACGAATACGTCCTCAAACTCCAGATCATCATCATC
TTCGTCCTCATCCTGGCTGAAACTGTAGTTGTGGTTGTCATAGGAGGCGGACGGAACTGAGGGCTTCGCAGGAGCAGGAGGAGCTGGTGGACTGTGGAACTGCTATTTCGT
TCACTGCTGCGCTGGGAACCCGGAGCGGGGAAGGGAACTCGTAGCTGGTTTCACCCAGCTCGGCGCCACTGGAGGCCTCCACTATGTTCCACATTTCTCGCTCGAGATCTG
TGGTCACCGAAGCCACGGCGGCCACCGTTTGGTTGATGTTGTTGAAACAGGGCGACACCACTCCGCACTTGGGCAGGAGCTGGGAGTTCTTCAGGGCCTCCTGCTGCTGCCT
CAACCAAATGGGTGAATCTGCAGGGAATGTTACGAGAGTTATAGGGATTTTGCCACAGTTCCAAGGAGACAACGCGCACTCACACTCACCCACATAGAGGCGCGGCTGCGTG
GAGCGCATATTGGAGTCCATGGACTGCAGGCGATGCTGACTGGCCACCGAGAAGCGGTGGATTGCCTCCGCCTGCTGATCAATGTCCACGCCCATCATCTTCAGGAGTAAGT
TCTTCAGCGCGGCCCTGAAGTCCTTGAGGTGATAGGCATACAGGACCGGATTTACCGCCGAGTTCAGATGGGACAGAATGATGCAGAAGAGCGTCAGCTTGGGATGCACATA
GCAGTCGGGACAGAAGGCCTTGATGCAGTTGATCGTATAGAGCGGTATCCAGCAGATCATAAAGAACAGCACGATAATGGACAGATTCTGGGTGGCCTTGACGTCCCGCTTC
CTTGCGGCGCCCAAGACCCGCAGCATGGTTCCCGTGTGTCCACCTCTTCCGGGTGTCGTCACCTGCACCACGGCCGCCGAGGAGCGACGACTGAGATCGGAGGCGGGGTTCA
TCGTAACAATCTGACGGACCTTTAAAGTTCAATTGAAAACGGATAGTTAATATTCCAAACTCTATGGAATATATAAAAAACTTTACTGGCGAAAATCTATCTACCACTCACC
TGTTTGATGATGACCCGGTAGATGTGCGTGTAGAAGGCCAGCATCAGCAGAGCCGGAGTGATAATGGTGGCAAAATAGAGGAAGACGAGGTAGTTGTAGTCCATCACCTCCA
CGAAGAGACACTCCTGGTTGTGATTGACATCGGCGTGCCAGCCGAAAAGCGGCAGGAAGCCCACTATCGTCCCGGCCACCCAGCACATGGAGATGATGACTGCAACGGAAGA
GGAACATTGGTAATCAGAGTGCCCTGAACTGAAGCCCTTGAAGTGGAGCTCCCCCAGACGTCATTTGAGCAGATTGCAGGAATCCCTGCCGTCTTCGCCACCAAGCACCATA
AAAACTTTTATCGAATTGCGGGAAACTAAACAAAACTGGCAGGCAGGGTGAGTGCACTTCCCAAATGACGCAAATAATTACCAATTTTCCGGACGTAGCATATTGGAAAGTA
ATCTTGCCTTGAGCGGCCATTTATAAAGCACTATATAGTAGGTGATATATAGTGCTTAGAAAATTATATATTTATGAGATATTTTTAGCTAGAGGCCAATTATTCTGATTCC
AAACAGTTGCTATAAACATCTAATATGAAAATTCTTAGAATCAATTTGAATCTAGTTAGGCTAGTAAGACTATTATGTTACGGATTTTAAGGATGTTTATTACTGTGTACTG
TTCTCATTGCGTCCATTCGTGGTAGTTCAAATGAGGCGTCTTTTGTTTAACTCAAGTCGCAACTCAAACATGGTCGATGGCCGGAATGAAACACAATGTGCTGTATATTAGC
CATCCAGAGTCATTAGTATCACGGCAACAAAGGCACAATACTAAACAATGGACAACTTAAAGTCGAAATAGTTTGACAAATATCCTGACAAGTAGGTGAAACCAGAAAACTG
GTCCTTTAAAGTGGGTTTTAGCATACCATTTTGTACCTACTGATTTTATACAATACATTGTATATACCCCCATATATACTTACATATCGCCGTGCGGGTGCGGACATTCCTT
GAGTAGGCCATCGGATATAGGATGGCCCAGTATCGATCCACGGACACGGCCACCAGACAGAAGATGGAGATGGTACACAGCACCACGAGCAGGGAGACTGTGAAGAGGCAGG
CATGAAGGTTTCTGGGTAGTCCCATGGAGGCCAGGATGGCGAAGGGAATACCCAATGCGCCCACTAGCAAATCGGCCATGGCCAGGGATACTATGTAGTAGTTGGTGCGGCG
TCGCAACTTTCTTTCCCGGCGGAAGACGATGATCACCAGCACGTTGCCGATGATGCTGACGATGGCCACCAGGACCTCGAAGACGGTGTAGGGTATGTTCAGCTCGGAACTG
GGGCTGTCCGAGTCCTTGGCATCCTTGCTCGTGGTGGCCGCGTGCAGCGGTAGCAGAGGTCCCTCGAAGGAGAAATCGGTGATCGAGAAGTAGCGAAACGCGGACATGGTGG
GCGGTGTTGCGGGTATTTTACGGGTTCCACGGGGTTCCCTGCCCCGATCGAGTATTTAAGGCTAGTGCCGAGCTAATGAAAGGGCGGAACAGGTGTGGGCTTCAATTACAGT
GCCCGCTGGCCTGTGAACTCCCTAATTGACTTGAGGTGGCCGTGGCAGTAGCAGCGATAGTAATAGCACATGCTGCGCATTGCGTATACGCCATGCCCGCCGCGCTGCGGTT
GCTCCTTTCCCGCTTTCTCGCTTTTCCGATTCGCTTTTTTCGCTCTCGCTCTTTATTTTCCTTTCCGTTGATGTATTTGTTTTTCTGCTGTCGACGTACCCTATTAAGCTGG
AAAAAAAAGGGAAAACTTAATCATCGTTCGGCACGTTTGTTACACACATGCAGCAACTTAATTGTAGCAATTTTTTAGTGTAATAAATTGATAAATATTTGCCAAAAAAATAC
ATGTCCATATTGATAACGTGTGCGGCGAGCAAAACAACAAAGGCCGTATCACAACTATGCAATACCCTAGATAATCGGCCAGTACTTGAGTATATTATTTGAAAGGGGCCAT
CAGTTGCAGGTCTTCCTTAAAGCCAGCAGCTCAAGGATCACAGTGAAATATGTTGGAACTTCGATGGCTTTTACGATGATGGTGTTTCTGGTTTGGCTTATGGGGTATATAT
ACCCGCAACTCTGCGAGTACACACTGCCACACTATCTGATTCGCGTGCGGCAAATACAAAGCTCGAAAAGCAAACAACTCCCGACTAGCAACAAAATCAATTGCCAGACAAT
CGAACCGGGCGCTGCAGAACATAATAAAAGTGTATAAAAAGGACGGAACAGCGAGAACAACGCAGAAACAGACGTCCCATAAATAAAAGAGCATCCTACACAGAGAAAAAAG
AATTAGGCAGGGGTTTACGAAAATAATTCATATTCAATTCAAAAGTATTCGCTTTTTTTGTAATTATTAAGATAGATGATGTATTGCTCCTAAAAACTC
(SEQ ID NO: 109)

Exon: 4251..3780
Exon: 3011..2800
Exon: 2707..2212
Exon: 2145..1001
Start ATG: 4251 (Reverse strand: CAT)

Transcript No. : CT27563
ATGTCCGCGTTTCGCTACTTCTCGATCACCGATTTCTCCTTCGAGGGACCTCTGCTACCGCTGCACGCGGCCACCACGAGCAAGGATGCCAAGGACTCGGACAGCCCCAGTT
CCGAGCTGAACATACCCTACACCGTCTTCGAGGTCCTGGTGGCCATCGTCAGCATCATCGGCAACGTGCTGGTGATCATCGTCTTCCGCCGGGAAAGAAAGTTGCGACGCCG
CACCAACTACTACATAGTATCCCTGGCCATGGCCGATTTGCTAGTGGGCGCATTGGGTATTCCCTTCGCCATCCTGGCCTCCATGGGACTACCCAGAAACCTTCATGCCTGC
CTCTTCACAGTCTCCCTGCTCGTGGTGCTGTGTACCATCTCCATCTTCTGTCTGGTGGCCGTGTCCGTGGATCGATACTGGGCCATCCTATATCCGATGGCCTACTCAAGGA
ATGTCCGCACCCGCACGGCGATATTCATCATCTCCATGTGCTGGGTGGCCGGGACGATAGTGGGCTTCCTGCCGCTTTTCGGCTGGCACGCCGATGTCAATCACAACCAGGA
GTGTCTCTTCGTGGAGGTGATGGACTACAACTACCTCGTCTTCCTCTATTTGCCACCATTATCACTCCGGCTCTGCTGATGCTGGCCTTCTACACGCACATCTACCGGGTC
ATCATCAAACAGGTCCGTCAGATTGTTACGATGAACCCCGCCTCCGATCTCAGTCGTCGCTCCTCGGCGGCCGTGGTGCAGGTGACGACACCCGGAAGAGGTGGACACACGG
GAACCATGCTGCGGGTCTTGGGCGCCGCAAGGAAGCGGGACGTCAAGGCCACCCAGAATCTGTCCATTATCGTGCTGTTCTTTATGATCTGCTGGATACCGCTCTATACGAT
CAACTGCATCAAGGCCTTCTGTCCCGACTGCTATGTGCATCCCAAGCTGACGCTCTTCTGCATCATTCTGTCCCATCTGAACTCGGCGGTAAATCCGGTCCTGTATGCCTAT
CACCTCAAGGACTTCAGGGCCGCGCTGAAGAACTTACTCCTGAAGATGATGGGCGTGGACATTGATCAGCAGGCGGAGGCAATCCACCGCTTCTCGGTGGCCAGTCAGCATC
GCCTGCAGTCCATGGACTCCAATATGCGCTCCACGCAGCCGCGCCTCTATGTGGGTGAGTATTCACCCATTTGGTTGAGGCAGCAGCAGGAGGCCCTGAAGAACTCCCAGCT
CCTGCCCAAGTGCGGAGTGGTGTCGCCCTGTTTCAACAACATCAACCAAACGGTGGCCGCCGTGGCTTCGGTGACCACAGATCTCGAGCGAGAAATGTGGAACATAGTGGAG
GCCTCCAGTGGCGCCGAGCTGGGTGAAACCAGCTACGAGTTCCCTTCCCCCGCTCCGGGTTCCCAGCGCAGCAGTGAACGAAATAGCAGTTCCACAGTCCCACCAGCTCCTC
CTGCTCCTGCGAAGCCCTCAGTTCCGTCCGCCTCCTATGACAACCACAACTACAGTTTCAGCCAGGATGAGGACGAAGATGATGATGATCTGGAGTTTGAGGACGTATTCGT
GCCGGCCAGCTCTGTTCCAAATCCCGTTCAGCCTGGCATAGATCCCGTGGAACTGCGTCGCTCCCTGGCTTTGGTCATGAGGGAGAAATTGCGATCGGATGACACGGACTCC

AGGCCAATGGGCAACAATCAGGATCTTCCCATAGATGAACAGTCCAGGGAGAGACCGCTCTCCACTCAAACATCTCCCACAAATGGCCCACTTCCGGCTCTTCTGAGGGCCA
AACTGCTTGCTGGCAACTCCAATAGCGCTCACTGCCTGCCAGGATCCACGGCGAGTCCTGCTCCCCAGGAGCAATCCGGTATCTTTGTGATCGATAGTGAGGCGAGTCCCGG
CTCAAATGGGCACAAGCCTAAGTATCGAAAGGGCACGGCATTCACTCGGAGTTCGCTGAAGAAGAGCCGATCCTGCAACTGTAGCTCCATCGCTAAGGGACGAGGGGTCCAC
GACGAGCCCAGCAGTAATCTCTGCAGGGATCAGGAGTCCTCTGTACTTCCACAGCATCCGCAGCCAGCCAACCATCCCACAGAGAACTTTTTAGTCCGCTGAGATCGGTGG
GCAGTTTCATGCAGCATTCCAACTTGTTTCACTTCCTCCAGCCTCACGCCGCCCGTCCCACCTCATCGACGGCGTCGTCCACGGCCTCCACGCCCACTCCATCACCACCGCC
CATGGGCCAGGCTCAGGAGGAATCGGTTCCAGTGGGTCTGACCACCTCCTCGCCGTCCCTTCTGGCGACCAGCGCGGAGAGTTGA
(SEQ ID NO: 110)

Start ATG: 1 (Reverse strand: CAT)

MSAFRYFSITDFSFEGPLLPLHAATTSKDAKDSDSPSSELNIPYTVFEVLVAIVSIIGNVLVIIVFRRERKLRRRTNYYIVSLAMADLLVGALGIPFAILASMGLPRNLHAC
LFTVSLLVVLCTISIFCLVAVSVDRYWAILYPMAYSRNVRTRTAIFIISMCWVAGTIVGFLPLFGWHADVNHNQECLFVEVMDYNYLVFLYFATIITPALLMLAFYTHIYRV
IIKQVRQIVTMNPASDLSRRSSAAVVQVTTPGRGGHTGTMLRVLGAARKRDVKATQNLSIIVLFFMICWIPLYTINCIKAFCPDCYVHPKLTLFCIILSHLNSAVNPVLYAY
HLKDFRAALKNLLLKMMGVDIDQQAEAIHRFSVASQHRLQSMDSNMRSTQPRLYVGEYSPIWLRQQQEALKNSQLLPKCGVVSPCFNNINQTVAAVASVTTDLEREMWNIVE
ASSGAELGETSYEFPSPAPGSQRSSERNSSSTVPPAPPAPAKPSVPSASYDNHNYSFSQDEDEDDDDLEFEDVFVPASSVPNPVQPGIDPVELRRSLALVMREKLRSDDTDS
RPMGNNQDLPIDEQSRERPLSTQTSPTNGPLPALLRAKLLAGNSNSAHCLPGSTASPAPQEQSGIFVIDSEASPGSNGHKPKYRKGTAFTRSSLKKSRSCNCSSIAKGRGVH
DEPSSNLCRDQESSVLPQHPQPANHPTENFFSPLRSVGSFMQHSNLFHFLQPHAARPTSSTASSTASTPTPSPPPMGQAQEESVPVGLTTSSPSLLATSAES*
(SEQ ID NO: 111)

Name: Adenosine receptor like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384520
GAGCCAATTCTTTTTTGGCTTTACGTGACGTTGGCAGAACAAATCGTTTTTGTTTTAGCCATCTGACTTGTTTGCTCCTTTATGGCCCGCTGCGCTTTGGACTTTGACGCAT
TTGCGTCGCTTTAATGCGATTGCTGGTTGCTCATTACGGCCAATCCTTACTAGTACTTGCACTTTCTGTGAATTCGTCTGGGGTTTCCATTTGGCTGATTGCATTTTCATAA
CAATTTGGAGGATTACTAATGCGCCACGTTGTCCCTGCATTTGCATGACCTTGAACCAGTAGCCACCTCATTAGAATTTCATTTGTAGTCAACAAGGCGCAGACGACAGGGC
CACAATCGATAAAAACCCAAATGTATTTTATGGCCAAATTACATTTTTACTAGAATGAATCCATAGAAAGATGCGGAAAATAAGCGGGCGCGGAATATGCTAAGCTTTATTT
GCAGCCATGGGAATAATCTTAATACAGCAAATTCATCCTTTAAGGTATCTCTAGAAGTAAGTTTAGTAAGCGAACTTTGTCCGTCATTTCCATTGAGTTGAATGCTACATTT
TCGAAATCTTATTTTAAATGCGATTCAATACAATATACAGAATCTACTATGTTTAAGATCGCAACATTTTTTTTACACTTTCTTAACTAAACACTTTAAAAGCACATCTGT
GTTACACTTTGGTATGCATGTGAAACATTTGGTTAACTAAAACATATATCTACGCAATTTAAAAGGCGATCTATTAAATTTCAAGGGGCAAGCTTTTCGTTTTTAAATTTAA
TTCCTATATTTAAATGGTGGGCGTTCACTAGTGATTGATTTATGCTTACCATTATCGATAATATTATTAAGTTATATCCGCCTCAGGCTTGTTTAAGAGTAAATCCCCCGAG
CTCTGACCAAAGTTTAATTTTAAGAGCGCCAGTATGCTGAATAGTATGAGTCAAAAGATTGAATATTCTTAGACTGAATAATTCTGAATACTTGAAAGACTAGTTTAGGAGT
TTTTCCAGCCGTTGTCCGCAGCTGAATGAATGAAAAACTTAATGCTTAAGGGAGAGTGTGCGCTCGACTGTGATTAACTTTTAATTGCCAAATGCAAATGAATGCAAAAGCA
GTTTAAACAAAACGCTTTGGTAGCAGCAGCAATAGCAATAGACGACAAATAACTGGCTTAAACAGAATTCAAATAAATTAGCATCAGGAAGCCGCACGAGAGAGTGAGTTCG
AGTTTATCAGTTAGTTGACTTGGACACCGATCATGGCGTACTGAGCGGCGACTGGCGGCTTGTTCGTCGGATTCTGGATGTGGAGTACACATAGAGAGAGAGAGAGAGAGAG
AGAGAGCCAGTTAATATCGAACCTATCCATTGATTATAGGGGGCACTTGCCTGCATTATCGACGGTTTCGGCTGCTCCGCCGACTCAATGGAGAAGCGCTGCGTTTGACTGG
AATTGGTTAGTGCCCTCCTCAGGCGGCGCGATTCGATGTTGTTTCGCGAATTCAGCGAACCCTTCGATACCTTCTTGCCAAACAGAACGGCCTAAAGGAGTGGAAATCGGT
AAGGTAATTAGTTAAATAGAAAAATTCAGACATTCGAATTATTAATAAATATATTAATAAATATATATAAATATTTATAATACAGATAAATTTGAAATATAGTATTGCGCAT
ACTATAACCCCCACTTCCGTTTTACTCACCTTGAATGCCTCTCGGAACTTGTGGCTCATAATGTTATATAACAGCGGGTTGATGCAGGTGGACAAGTAGTAGAGGACACCGG
AGACATAGGTCATCACCGTGTAGACAAACTCGTGCTGATCCCGCAGTTTGGCCCCCCGTGCAGGGGCGTAGATGGCAATCAGTCGCTGGGCGTGGAAGGGGGCCCAGCAGAG
GAAGAAGCACACCACCACGGCCACTAAAGGGGAGGGGGGGCGGTAAAAGTCAGTGACGTGGGGATGAAAATGGAATGGCAATGGAAATGGGGAAGACACAGCGAGAGTCAC
AGATAAAGAGTCAACGATTATATTTTTTGGTAATCGCAATTGTCATCATAAATGGAGAGTGCCAAAATGACAAACGCAGTAAATTGGCTGCAGACGGGTTGATTAGCCCAGC
TTAGGGCATATGGTCAAACAGTTGGCATACACTGAGCAAAAGTTGAAGTAAAGCTTCTAAAATGGTTCTTCATTGCATATAAAGAAAGATTTAGATATCATATTTCAGAATT
TTTAATTTGCAAGACTTTTGTGAGAAGACTAAACCTTATTAAAAACGGAGTCTTCTACTTTTGTTTTTGTCAAAATTTCTTACTAATTTAATCATCTGAATACTATTAAGTC
ATGTTTACTTAATGGTTAGCGGTTTAGAGATCTGTCATTTAATTTGTGATACCCCCGTACTATTGCTGGTATTTTTCGCCAGTGTATATCGTGCTATAATTACCTAGCATCC
TGAGTACTCGCCGGGTGCCATAGTGATTGAGCCGACCTCTCACCGAGCTAAGTTGAGCCCCCGAGCCGCCCATCAAGCCCGCTGTCCCTGCTCCACTTCCTCCGCCGTTGAA
ACTCATAGCGGTACCGGATCCACCATAGCGATATAGGATCGTATCACTGGGCACACTCTTCAGCTGCTGCCGTCTGGCAACCGAGGCAGGACCCTCCACCAAAGTGGATCGA
TACAGGTGCACACCGATAAGTAGGTACAACACTAGGATAATAGACATCGGAGCCAGGAAGAATATGAACGTGGACAGCTGGAATGAGTGCTTCACTATGACCCGCACTATGC
CGCATTGCTCCACTCCCGAATAGTGCTCGATTCCAAATTGGGCAGCCTGCGGAATGGCCGTAACTATGGCCATAATCCAAACCAGGACGATGATGCGAATGGCGCGACTGAG
TTTACTCATGGCCTGGCCCAGAAACGGATGGCAAATGGCAATATACCGCTCCACCGTGAAGGCCGTAATCGTTAGCACCGTGGCATTCGCCGATGTCTCCGCCAACAGACCA
CGTCCGATGCAGATGTACTCCCCAAACACGTACGGGTACTTGGACCAGATGTAGGACACCTCCTGCGGAACGCCCGACAACAGGAGCAGGAAATCCGAGATGGCCAGCGAAA
AGAGGTAGTAATTCGTGGCCGTGTGCATTGAACGGTTCTTCTTAATCACAATGCAGGTACTTATGTTGCCCACTACACCGGTTATGAAAATGAGGGAGTAGACCACCGTTAC
GGGTATCACGATGGCCAGCGGATCGCGTGGCGGTCCAAGATCATGGCTCATATTGCCAGCGGACATGGCGCCGGCGGCCTGGAATTTTTCCCCTTTTGTTTTCACAAAATTT
CCACTAACAGCGCTAAATTCCACTGCCGTTTAACTGCTGCGGCTCCGGCTGCGACTCGATGTCCCCTCAACTATGCTAATTATAATAATTTATTGGAGGCGACCGGGCGGAG
GAGGGACAGGTCTGGACTCGCGCTTCGCAATGCTAGATACTCGATTCTAAGATTCAGATACTCGACTGGCACGCAAACGCACCAAATTACCGACGAGGTCGAGATCCGTATA
CCGTATTCAGTATTCAGTATTCCGTATTCCCGGCTGTTTTTCTAGGGGGCCCGACTTCTCTTCAATAAGTTATCGAGTTAATTTGATTTCGATTCGAATTCTGATTCCCGAG
AGCAGCCTGCAACCGCACTCCAGCGCGCTCCAAACTTATCTGATAAGCCGCGGCGACATTTCATTCAGATTTGTATCTGTATCTGAAAGCCAGCGACGAAGCGTCGCTGTTT
TTCCCTCACATGAACGCGGAAGGTGAGCGACATGGTTGTAATGACCGAGCGTCCAGCGAACGTGCCGGCTCAATTAAATCTTGAAAAGCGGCGATATCTTGATGGGACTAGC
CGATGTTAGAGACTAAGCACAATGGCTCCGAATACAACGCAATTTTTGGGTAATTAAAAAGGCTAATTAACAAAAACTTGTAATAATGTTTGCTTTGGCCCATTTATAAAAC
ACTTAGTTGGGAATTTTGACTTTGTGATGCACATAATCCTAAATAAACCACGATAATGTAAACAACACATCTTGGAATTAAAAGCCAAAATAAGATGCAATAATATTAGATG
ATGCAATAATACCAAATGACCGAAGGCCACCGTGCGTATGAGTAATGTCCAGCCAGCTTACTTGTTAGCTTATAAGTAAGAATATTGTTCTTTGTTTATCTTACTAAAGTAT
ATTGAAATGTAAACAAATAATAATAATTACAATTCAAATCTAATTTAGATTAATTAAA
(SEQ ID NO: 112)

Exon: 3314..2456
Exon: 1927..1710
Exon: 1547..1367
Exon: 1029..1001
Start ATG: 3314 (Reverse strand: CAT)

Transcript No. : CT27924
ATGTCCGCTGGCAATATGAGCCATGATCTTGGACCGCCACGCGATCCGCTGGCCATCGTGATACCCGTAACGGTGGTCTACTCCCTCATTTTCATAACCGGTGTAGTGGGCA
ACATAAGTACCTGCATTGTGATTAAGAAGAACCGTTCAATGCACACGGCCACGAATTACTACCTCTTTTCGCTGGCCATCTCGGATTTCCTGCTCCTGTTGTCGGGCGTTCC
GCAGGAGGTGTCCTACATCTGGTCCAAGTACCCGTACGTGTTTGGGGAGTACATCTGCATCGGACGTGGTCTGTTGGCGGAGACATCGGCGAATGCCACGGTGCTAACGATT
ACGGCCTTCACGGTGGAGCGGTATATTGCCATTTGCCATCCGTTTCTGGGCCAGGCCATGAGTAAACTCAGTCGCGCCATTCGCATCATCGTCCTGGTTTGGATTATGGCCA

TAGTTACGGCCATTCCGCAGGCTGCCCAATTTGGAATCGAGCACTATTCGGGAGTGGAGCAATGCGGCATAGTGCGGGTCATAGTGAAGCACTCATTCCAGCTGTCCACGTT
CATATTCTTCCTGGCTCCGATGTCTATTATCCTAGTGTTGTACCTACTTATCGGTGTGCACCTGTATCGATCCACTTTGGTGGAGGGTCCTGCCTCGGTTGCCAGACGGCAG
CAGCTGAAGAGTGTGCCCAGTGATACGATCCTATATCGCTATGGTGGATCCGGTACCGCTATGAGTTTCAACGGCGGAGGAAGTGGAGCAGGGACAGCGGGCTTGATGGGCG
GCTCGGGGGCTCAACTTAGCTCGGTGAGAGGTCGGCTCAATCACTATGGCACCCGGCGAGTACTCAGGATGCTAGTGGCCGTGGTGGTGTGCTTCTTCCTCTGCTGGGCCCC
CTTCCACGCCCAGCGACTGATTGCCATCTACGCCCCTGCACGGGGGGCCAAACTGCGGGATCAGCACGAGTTTGTCTACACGGTGATGACCTATGTCTCCGGTGTCCTCTAC
TACTTGTCCACCTGCATCAACCCGCTGTTATATAACATTATGAGCCACAAGTTCCGAGAGGCATTCAAGGCCGTTCTGTTTGGCAAGAAGGTATCGAAGGGTTCGCTGAATT
CGCGAAACAACATCGAATCGCGCCGCCTGAGGAGGGCACTAACCAATTCCAGTCAAACGCAGCGCTTCTCCATTGAGTCGGCGGAGCAGCCGAAACCGTCGATAATGCAGGC
AAGTGCCCCCTATAATCAATGGATAGCTGCGGACAACGGCTGGAAAAACTCCTAA
(SEQ ID NO: 113)

Start ATG: 1 (Reverse strand: CAT)

MSAGNMSHDLGPPRDPLAIVIPVTVVYSLIFITGVVGNISTCIVIKKNRSMHTATNYYLFSLAISDFLLLLSGVPQEVSYIWSKYPYVFGEYICIGRGLLAETSANATVLTI
TAFTVERYIAICHPFLGQAMSKLSRAIRIIVLVWIMAIVTAIPQAAQFGIEHYSGVEQCGIVRVIVKHSFQLSTFIFFLAPMSIILVLYLLIGVHLYRSTLVEGPASVARRQ
QLKSVPSDTILYRYGGSGTAMSFNGGGSGAGTAGLMGGSGAQLSSVRGRLNHYGTRRVLRMLVAVVVCFFLCWAPFHAQRLIAIYAPARGAKLRDQHEFVYTVMTYVSGVLY
YLSTCINPLLYNIMSHKFREAFKAVLFGKKVSKGSLNSRNNIESRRLRRALTNSSQTQRFSIESAEQPKPSIMQASAPYNQWIAADNGWKNS*
(SEQ ID NO: 114)

Name: Growth hormone secretagogue /neurotensin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384297
TTAAGCCACAAAATGATTGAAAGGAAATGTTGTGGGGTAAAAGTAAAATTCTCAGATCGATGAGTACAATGTACTATGTACTTCCCTCGGCAGCTCCTGCTGTGTTCCTTCA
GCCGTTTGTGTCTGTTTGTCCCTAAAGTAATTAAAAGTAAGTAAAGCAAAGTCACCCCCTCGTCGGTGGGCGTGGCTGTCCAGTGATTTGCCACCCGCAACGATTTTTTTT
CCAGTACTTTCCTTTTTTTCTTTTTTTTGCGTTTATGCAAATGGCCGCTGACCAAAAGAATTAATGCGTCCCGAAAGTGAGTTGAATTGGAATGAGCCGCAATTCTCGGAGT
CCAAAAACTGCGGGCAGCCTGTCAGCCGTCTGTTTTCACTTCTTTTGCGGGCGGGACGAAACATCAACGAATGTCAATGAATTTACAGGGGAATGCTATAAAAAGCCCGACC
AAACAAAAAATGTAAAGCCCGTAATTCCCCATTGTCTCAGTGCGTTGCTAATTTGGCCATATCTTTAAATAGCCAGCCGAACGGAAAGGAATCATTTGTAATTTAAAGTCCA
ATTATTGTGCCATTTGAGGGGAAAACTGTGGCATGCCAGGGAATGGAATGCCTCGTAAGCCACATCGCTGCACTTCATTAAATTGACGAGACAAAAGATCTCTACACATACA
ATAATGGAAAATATTTGTTCCACAATAAAAGGGGTTGAGGAGGCACTTTGTGTTAATGCCCAACAGCTGCAAATCTCGGTCTGCGGCGGCCTTAAAAATAGAGAAATGAAGA
TGGACGGTCCGATGATGATCCGCTTGGCTGAAATTATTCTGTAAATTCTACCTATTTATATCGTAAATATAGATTTTGATGCTGATGCCGATGCTGCGCGAGTAAAACCAAA
TTCGGCCCACGCGGCGTATGCGCGTTTTTCGCACTTAACTCATTAAAACGTTTGCCGCTGGCCCATTGAAATTTAATGGGAAAACACGCGCTGTTTTTTTTTTTTTAGAAAT
TCGCCTCAATAGATTAAGCTCGTAAATTATTTAATTGCATTCTTGATTTTTGATTTGAAATCAAAACATTCCCTCATTTAAAAACTCATTAGGAATAGCAATGGAAACGTGG
TTGAGCGGATGTTTATATTGTAGAAAAAAAAAATATATATATATCAAATATATGAAAAAGGTTTTGAAGTTTGTCAGTCAATAAATCAATATAAAGCCCCACTTAAAATTAC
GACACAAAAGTTTCCATCAAACGAAATGCAAACGAAAACAATGCTCTGAAAATATTTCATCGATATATTGACCGAGCACTCAGATGTCGTCACTTGAAGGACCCTTCTCGTC
CAGCCCCCGAGTGCACCAACCCCACTCACCTTGTAGAAGGCCTTCCGGAAATTCTCGGAGAGGAAGGCGTAGAGCAGCGGATTGATACACGAGCTGCTGTAGGCCAGAGTCT
GGGCGGTGACCTGGATGACTAGCTTGGTGAGGGTGTTCGTCTCGATGACATCCAGTGACTTGAGCAGCAGGATGAGCTGTTCAATCAAACGAAAGAATATGTTTGATTATTT
AAAATATATGTGGGTTTCATATGTATTATTGCGTCTCACCTGGACAGGCAGCCAGAGCGAGGCGAAGGCGATGACCACCACGACGACGAGTCGGGTGACCCGCTTGCGACCG
CGCTGCGACTCCTTGGACATGCGGACGCCGGTTCCCTGGCGCCAGAGCCGCATGATCATGCGCATGTAGAGACCGCTGATGATCATCAGGGGCAGCAGGTAGGAGCTGATGA
AGAAGGTGACCTGGTAGGTGCGCGGACCAAGGAAGTCGTTCGTCGTGAAGGTGCACATGCCGTAGGTGATGTTCTTCTTTGCATCGTAATCCACCTGCGGATATCGATGAGA
ATGTGGTTGTGGGAACTATGAAGAAACGTGTGCGGTGTGGGAATCGAATATTGCGCATACGTCATGTTGGCGGTGACAGTTGCATGGCAGCCAAGCAACGTCAAACGGCCAA
CAGCCAGACATCGTAAATAAAACTAAACTGTCATACGAATAGTTCGGACGTCGACAACTAACTCGCAATCTCACTTTGTTTCTCTGTCGCTGACATTTGGCTGACAAATTGT
TCAGCCGGAGTTTTGGATTTGGATAAAAGCGCAATGGGTATGGGCAATCGGGCTTTCGAGGCTTGTTGTTTGGGGGATGGAGAATCTATTGGTCTGTTGGCTAGACGAGACT
GGTTTGCTGACGCTTAGGAGGGTCGTAAACTGCCCGGTGAATCGTAAAAATATTCGCAGCCTTTTGTTTACATTTTTGTTTACAATCTCCGCCCCTTCACCGCGCTTTCAAT
CGATGAATGCGTGTGGAAAATTGCATTTAAAAAATCTTCCCATTAATGCCGGCATGAATTGCTTACAAAAATGTTTGTTTATGGCCCGCCCTTGTTTTCGTTACACTGAGCG
AGCGATGTTGGCCCACAAATTGAGTTCGAACTCTGAATCTGCGCAGCGTCAAGTGGCCAAGGCCACAGAAGTAGACAAATCATGGCCAGACCCTCCGCTGATTTATGGGTTC
ATTTCGCTGGCATGTTCTCCGCAAAACCCAACGACTTCATAGACTTGACAAGCCATTAATGCCCTTTCCTCGTTCGCGAGAAATCAGCCAACAGCTGTCAATCAATGCCATG
ACCTGGAGATCGCATATCTAGTTACGCAGTCGGACATATGTGCATTTATGATTATTATTTGATATCGATTTCTAATAAAAAGATAATGGCACACGCTCTCACAGGCTCATTT
GATATGCATATACCAAGCATGTAGCTGGGCAATGTAATATGGAAATAATTAATAGCAAACTGTTAATTAACATGTTGATCAAACATGACTTAATTATGCCTTTTTATCGCCT
GTTCCTTTGTATTTGCACACGCAGACACCAGATAACAAAGCAAAGCACACATATGAGCATTCAATGTGGGCTCTATTTTTCCAGCTTAATCCTTAGACAAAGCAAAGCAAAC
ACATAAACAAACAATAATGCCACCCATTGGAGCACTCTCCCCAATGCGAACAAAAACCGAACGAGACAAAAATAACAATATAACGAAGCGTGAAATGCCATTTAATAAAATA
TTAGGTTACAATAACAACACCGGGAGCAACAGTGGGCCATTGCCATCTCACAGATAAAAAAAGTTCATCATCCGTGTGAAGATGAGATTCGTTTAATTAACCGCGACTGTCG
CTTTTGGCCAAGGAAAAGGGGAAACACAACGGCACAAACAAAAAGTTGGTCATGTGTGATGTGGGAATCGCTTTATTGGCACTGGGAAATCAGTGGAAAAAGCCCGCGGGCC
ATTTATGAGTATCAACAGAAATTGGAACCAATGTCCGACCATAATTACGTGTAGTGATACTCGCAATGAAAAAATATGATAGATATTCAGTATTAAATAGCACTCAACAAT
TGTTTGAACTGTACCGGAAAATTGAATATTAGTTAATAATCCCAGTGTTTTCTTGCGAATAAAAACTTAAGGTTAAAGTCTGCAAATACAAATACATTGTATATCCTCAAAA
TCGCCGGTAAATTCCCCTTTTCATTTTTGATATCACCCGCTGATCTCGAGCAAAACAATTCCGATCTTCGTCGATTATTCCGCAATTAAATTGTTTCTCGCACTTTTCGCAT
CGATCATAAAAGCCTAATGTCCAAGATGAGCATTTCAAGGTCGCCACCCGCGAAGACATTTATCATACGCAGTGTGCGCCCAAAGTTCGCAAGTGCAAAGTTTCCATAATGG
CCGCTTGTTACGACTTACGACTGCCCATGATTTACGCGCTTTTGGATCAACAGTTGAATTTGAAATTTTGCAAGCCACACGGAGAGGAGGTGGGCAGATATGGGCAGATACC
GCCCAGTTGGCAAGTTGTTTCTCGCTGGCTATCTGGTCTATAATTAAAAATTTATGAATTCAAAATAGCAAATTAGTAACAAAAACAAATCGGGCGCCGCGCATCGCGAAATT
TTCAATATAAACACAAAACAGAAAATAAGATAAAATGGGGATACTTTAATAGGGGCATATAGTAAAAACAAATGAGTTGTATCAGGATCAGTCGATTGGACGATCGCTGGTT
TGTGCGGGGTCTTCGCTTTCTGGATTACGATCTGGCAGAAGGCAGACCATCTATCTAGGTATCTATAGCTACATTAGAACCAGGCAGTCGGCACCATTGACGTGCAGCGAGA
GGAATCTAACGAGGTGACGATGATGGAAAGCAATAAATTTGCCTGGCCCGTGTTCTTCACAGCAAATTCAGGCACAGCAGGCCCACTTATAGCATTCTACCCACTCCTCCCA
CTCTTCCACTCCCACTTACCACCACGTCGTGGGTGAAGGCCACTGGCACCGAAACGACCAGCACCACGATCCACAGAGTCACGATGGCAATCAGGGTAATGTTCTCCGTCCT
CATCATCCGCGAGCGAATGGGATGAACCACCGCCAGGAACCGATCGATGGACATCAGCACCAGCGTGTAAATGGAGGCGAAGGCGGTCACCACAATCAGGTACTGGACACTG
CGGCACCAGAACCTTCCATATGGCCAGTAGTACACCATGTAATCGGTGGCCGTGAAGGGAATGCAGAGGATTACGAACATCAGATCAGCGGCAGCCAGATTGACAATCATCA
GGTTGGTGGTGGAGCGCATGTTGTTGTTGAAGACCACCACCAGGATGACCAGCAGGTTGCCGAAGAATCCCGTGATGGCGATGACGCCGAAGAAGAATCCCACTATCCACGG
CAAAGTGCCATTGATGGCCAGCCACTCTTCGTCGGTGAAGAATGAGGTGATATTCTCCTCATTTCTGGTTGCATTTAGGCTAATATTAGCCAGCATTGTGGTGTTCTCCATG
TTTTCTCAGCTTTAGTGCACAACTTCCGGTTCACTTTGTACTGGCTATTGGCGATTTCGCTTGATTGTCACGGGACATTTTCAATTTTGGGCTCTTTTTCGCTCACAGTTC
ATCGAACGCAAATAACAAATTTCCAATTAGGAGAGTTTCACAAAGCAGCCGCCACAAAAATTCAATTATCATGCGGAATACTACGACCGTTGAGCTTCGGAGACCTCAAAAA
CGTTTCGCATACGCCGTGTGGTATCGAACTCGACCGGAATTGTTCGCAACGTGGCGGGGGCAAAAGCAATTTGGTCCTTCCACGATAACGGTAATTGTAGGTGGGTAGCACA
GGTAAGACAACAATAAAGGACAGCACACAAGGATCACGGTGAGGCGAAAGGCAAGAGAGTGAGTTTAAAGGTGAGGGCACGTGCAAACGACACCGACACAGCCAAACTGTGC
CGAGTCACACCCTGACGACAAGGACACCGGGGTGTCCTTAGGTTCGAGGACCACGTTGCTGCACCTACACCATAAATTCAATATGATGTCTTCCTTTCGCTGCGTTTCTGTT
TACATTTCTCGTTTGCTTTTTTATTGACAAACGAACGGCATAAATTTTGGGGCTATAAGGCTATGGCCAAAAAACCACGTTCCTATTGTTGCTACCGACTGCTGCGCCGCGG
TACGCTATATATTTTCGATTTTATACTGCTGCTTTTGGCCAATTCGATAGTCCACACTGCTCGCGTCCGTATCCAGCGCCCAACTGATTGCTGGCTGAGCTTGTTGAGGAAT
ACTTCGATGCCTCGGCGGCGGCAAAACGCCGGCAGCGACGCCAGCCATCCCAGCAAGAATAATGGCAACAATAGCAACAACCAGCGCAGCTGGCATCTGTGCGTGTGATGGT
TTTGGGTGGCCCGCCGGAGCCACCTCTTCCTCCTCATCCTCCTCATCCGATCCCAATCCCCATCAGCTGGGCTGGAATGTTTGCTCTTGTTGCTCCCACCGCTA

(SEQ ID NO: 115)

Exon: 4927..4388
Exon: 1886..1608
Exon: 1532..1374
Exon: 1012..1001
Start ATG: 4927 (Reverse strand: CAT)

Transcript No. : CT28187
ATGGAGAACACCACAATGCTGGCTAATATTAGCCTAAATGCAACCAGAAATGAGGAGAATATCACCTCATTCTTCACCGACGAAGAGTGGCTGGCCATCAATGGCACTTTGC
CGTGGATAGTGGGATTCTTCTTCGGCGTCATCGCCATCACGGGATTCTTCGGCAACCTGCTGGTCATCCTGGTGGTGGTCTTCAACAACAACATGCGCTCCACCACCAACCT
GATGATTGTCAATCTGGCTGCCGCTGATCTGATGTTCGTAATCCTCTGCATTCCCTTCACGGCCACCGATTACATGGTGTACTACTGGCCATATGGAAGGTTCTGGTGCCGC
AGTGTCCAGTACCTGATTGTGGTGACCGCCTTCGCCTCCATTTACACGCTGGTGCTGATGTCCATCGATCGGTTCCTGGCGGTGGTTCATCCCATTCGCTCGCGGATGATGA
GGACGGAGAACATTACCCTGATTGCCATCGTGACTCTGTGGATCGTGGTGCTGGTCGTTTCGGTGCCAGTGGCCTTCACCCACGACGTGGTGGTGGATTACGATGCAAAGAA
GAACATCACCTACGGCATGTGCACCTTCACGACGAACGACTTCCTTGGTCCGCGCACCTACCAGGTCACCTTCTTCATCAGCTCCTACCTGCTGCCCCTGATGATCATCAGC
GGTCTCTACATGCGCATGATCATGCGGCTCTGGCGCCAGGGAACCGGCGTCCGCATGTCCAAGGAGTCGCAGCGCGGTCGCAAGCGGGTCACCCGACTCGTCGTCGTGGTGG
TCATCGCCTTCGCCTCGCTCTGGCTGCCTGTCCAGCTCATCCTGCTGCTCAAGTCACTGGATGTCATCGAGACGAACACCCTCACCAAGCTAGTCATCCAGGTCACCGCCCA
GACTCTGGCCTACAGCAGCTCGTGTATCAATCCGCTGCTCTACGCCTTCCTCTCCGAGAATTTCCGGAAGGCCTTCTACAAGGCGAATTTCTAA
(SEQ ID NO: 116)

Start ATG: 1 (Reverse strand: CAT)

MENTTMLANISLNATRNEENITSFFTDEEWLAINGTLPWIVGFFFGVIAITGFFGNLLVILVVVFNNNMRSTTNLMIVNLAAADLMFVILCIPFTATDYMVYYWPYGRFWCR
SVQYLIVVTAFASIYTLVLMSIDRFLAVVHPIRSRMMRTENITLIAIVTLWIVVLVVSVPVAFTHDVVVDYDAKKNITYGMCTFTTNDFLGPRTYQVTFFISSYLLPLMIIS
GLYMRMIMRLWRQGTGVRMSKESQRGRKRVTRLVVVVVIAFASLWLPVQLILLLKSLDVIETNTLTKLVIQVTAQTLAYSSSCINPLLYAFLSENFRKAFYKANF*
(SEQ ID NO: 117)

Name: Allatostatin receptor type
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000012806866
ATCTTTTAAGTAAAAAGCGCTCAAGTCAAAACTTTTACTTTCATTTAAAAGTAAATTTTCTTGATGTTTCAAATTTATAGCAAAATGCCATAATTTCAGTTCGTTTGAAACA
AACGAAATGAAGCGGAAATTGCGAACCGAATTTATTCGTACAATTTATGGGCCACACATGAACAATAATAATTCATTCGGCTGCATGTAATTAAGAATAATAAAAATACTGG
TATATTTTGTTTGTTTGCTTAATGACTCGAAATGTTTATTCGTTTTCGTTTGGTAATTAAATACAATAATAATTACATTAATAAGCCTTTCTGTTTTTTAATGACAAATATT
GTACATTACATTTGCTTACATCCAGCTGTGTGCTTGTGTAGGATTTAATACATAATCGGATTGTTGTTAAAATACTATTTAAATAATTTAATTTAATGCTACTTAATTGATT
TCAGGTGCTAATTTAATTTGATTCATTTTTTGTTCAGCTAATAATACATTTGATTTGGTGAGTTTTGGTTTGAGTTTGATAAGGTCATTAACGAATCGATTATTTACAACGT
TTGGACTTCGTGCTTTAAGGATCGCAATTTGAAGATACCATTATTATAATACTGTACACGGCTCACGACCAGCATTCACACTCATTCAATATATATCCTTCTGCGAAACAAT
TATCAGCAATGGGACTTGTAATCACAATTCATTAAGACGAAATCACAGAGGCAAATGGTCATGTTCCAATGACACATTTAGTTGCGGGATATCTAAACACTGTAAAGGGGTT
TTACTAAAGCCTTACATACAATTACAATAAAATTTAAACATATTAGTGTAGTAGCTATACCCTAATAATAGTCATTCACTAAAACATTAACACACATACGTTAAATGAATGA
GCTATGACAGCCAATTTACTGGAATACACATCATATGTGAATGCCGACTGGCTGCGGAATGTGTCATTGGAACACGGCTATAACTTAGATCGCTCTAGTTTAAATTAAAGTG
GTTGCCACAAGGACTTGCCCGCATCCTTGTCCTTCGCCTGCGAGCCGGGCAGGTGGACTCGCGACACCTCGCTGAGTATGCAGGTGCCATCGCATTCGGTTTGGCTGCCAAA
TTCCTGTTCGCAAATGCAGCTGGAGCTCTCCCTGCTTAGCAGGGCCAACTGGACTCCGGCCCCGGATGCCAACTCCTCGCTGGGGGGCGTGGTTGATGGCAGCTCCATGAGT
GCCACGTCGTCCTCGCATTCGATTACGTTCGGATGCAGAGGTTTGAAGTTGTTTCGCCGCCACGGCGGCATTGCAACGCCAACACCTGGTGCCGAAACACCCGGAGTGGCCG
CAAAGGTAACCACGCTTTGATGCTGATGATGTTGGCCATTGACATTGTTGTTCTGCCCACTACTTCCGTTGTAAATGCCGCTGTTAGCACCGGATCCATGCACCCGCGGCAT
ATGCAAGCCCGGCTTCCCATTGTTGACACCACCACGCGCCGGACCAAAGAGATTACTCCGGATTCGCATCGAGGAGTTGGCGTAGGTTGACCTTATGGAGCTGGCGCGGGTG
TGCATCGAAAAGGTCCTTTCGTGGGCGTCCATGCTCGTCTTGAACTTGCAGAAACAGGCCGCAAATCGCTTGTTGAATTCCCGCTTAAATTTTTCCTGTGCAGAGAAAAATA
GACGACGGTTTAGGTTAATCCATCCCAATAACATATCACTAAAGTAAGCGAAAGTACTTTTATTCACTTAAATACAAATTAAAGTGCATTGCATAAATCATAAACCTCTTAG
CAAAGTTTATCATTTGCAAAAATACATTTTTTCGCTATCAAACTTTCTGGTTAATTATTATGTGAATTGGCAGAAAATTTTTCGTGAAAAAGCGAAAACTGCCCCGGCTACG
GCATCTTTCTGGCTAATGAAAGTGCTCTAGGCAAACCTCCATCTTCGCTGACAAATCCCATTTGTCTTCGCTGACAGTAAAATAGAATATCTGTTGACAAAGTTTTTCACGC
ACAGCCAGCAGGGGAAAAAGTTTACGAGTTTATTCGGAGATCTGGCTCTCAATCTATCTTTTACAGACAATTTGAGTGTCAAAGCAAACAGCCGACGGGTTTTTTCCCGCTT
TTATCGGTTTGCCACTTTGGCATTCTGAGGATTTTTGTGGATTTCGCTTCAGTTTTGCGGTTTCACAATTTCTGCGGCCCAGACGGTAGGTGTTTGCTTATTGGCTACAGTG
GAGCGCTGCTTTAATTAACTTTGACAAAGTATTTTTTTGTGTCAGCTATTTTAGGTGCGAGGCGAAAAGAAATTGGAAATTGTTTCAAAGTAAGGCAGAAGTTATTTTTGAT
TAAAAAAAAAAGTGACGCGTATGCTTATAAATCAGCGAGAAGGAAAATGGGAAAACATCAACAAAGCAATCGATTCGCTTACAATGAAATTAAGTTAGACATATTTCATTCTT
TTTTAAGCTTCTTTGATTTTTGAGACCATGAAAACTTCCGTAAAAGTTTCATTTGCTTTGCGTCTATGTTTGTGTTTCATTTGATGAGATAATAGAAAAAGCTAATTTTGGA
AGCACGTGCAAAGGCTTTTTTTCTTTTAGGTCTAAAAGTACGAAAATAATATAGAAAAAATTACTTTTTTTACTTTAACTTCATTATTAAAAGCAAGACGTTAAGTATGTAG
ATATTTTGACGAGACATCCAATTGTACGCAAGGCAAGTTTTCCCAAGATCCCCTGTCAATAACTGACTCACCCAAACTTAAGTTCCCCAAAACTAACCCAAGGCGACCTTGA
GTCCCTCATCAATCCCCACATTTCCCCGGCAAATTTACCACACATTTATAGACTTATGTCTTTTTAGTGCGTGTGCAAAGAACGTGTCCCTGAGCACGAACAAACATCCAAA
GGATGTCAATCAAATCCCAGCCAGACAAATGCAAATTTAACACTCCGTAAACAAATCATCATCATCATTGGAGCCCACACTTACATTGTAGATGCCATAAATAAAGGGATTG
TAGCAGCTATTGCTCATGGCCAGCCAATCGCAGCAAAACCAGACGATGCTAATGAAGTGGTAGTCGTTGATTTCCGGTATCGTGACATACAGAATATTATAGAGCTGCAGTG
GCAGCCAGCAGAGTCCAAAGATAATGACCACGATAATCAGCATTTTGATGACCTGCAAGCGTCGCAGAAAAAGGAAATTAAACCCAGTCCCGATTTGCATACCACATCAACG
TGCAAAACTCACCTTCTTCTTGTTTTTCAACAGCGTTATGTCCCGTGAATCCTGTGCGTTACCAGGAGCACGTGTGCCCCACAATCGTACCGCCATCTGGATGTAGACAAAG
CTGATGACACAGAATGGAACCAGATACTGCACAAAAACCAGGGTGTAGCGAAAGGATTGCAATTGATCATCGGATAGGTTCTTGTTCATGCAGAATGGCCGCGTCACATTGT
AGGTCTCATTGTTCTCTGTTCAAAGGATGAAATTCAGTTAAGGTCGCTGGCACTTAATCCATGCAAGTTGACGTATTGACTTACCGCGAAATCTTTCGGTCAACTCCTCCAC
ACGAAAGGCAATGGCAAAGGGCACCGCAAATAGCAGGGCCAGCATCCAAATTCCACCAATTATGAACTTCGATACGAACTTGGTGGGACGTGCCCTGCATGCGTAAGCCGAT
TTTTAATGATTATGATTGACGAATCGAATCGTTTCGAATGTGGGAGCCCACTCTCCCCCCCAGCGACTCACCTAAGTGGATTAATGATGGCCCTATGCCGATCGATTGCAAT
GGCGGTCAGCGTGAATACCGAGACATTTACACTCAGGGCCTGGACGAAGGGGCAGAAGCTGCACATGAACCACGGCAGGTTCCAACTCTGCAGCAGGGCAGCCTGGAACTAC
ACGGAATGAAAATATTGCAATTATTAGTTTAAAGTATTGCAGATTATCAGAAACAACTTGACTATACAACTATGCCTAACCTTTAGCCTGAGCATGATGTTTACAAATCACA
AATAATCCCTTTAAAATGTTCACTATATATATATTCATTTTAATTAAAGTGTAAGGAAATAGGATAGCTACAAGTGAGTCGTCCGCCTGAGGTTTGATATGATTTAATCAAC
CCAAAATTGCGACTCCATTCAAGTCACACGGCAAACACCATTAATGAACAGTGGAACATGTCGGGGGTGTTGTATTTTCGGCCAGTTTCAACATTTCGTCGACAAATAAACG
CTCGAAAGCCCAGCGGTGTTGTTTTGGACACTGATTTGCGCAGGAGTGAGTGAAAAAGTGTGTGTGTGATTACTGCGGATATTTGAAGACCGCGCTTTCTCCGTTCTCCACA
CAAAGTCGGCCGCAAATTGCAATTTCTCGCCGGCTTATTGATGACACTTACTGGCGGCTGTGACAGGACCAAAAATCTATTTGATTGCGGCATTAATATAAATTAAGTGCCC
AAAATGTGGCTGGCTTTGGGTCAAGTTCAAGGTTCAACGAAGTGTACAGATTGAGCTTTATTAGTATGAATTCTGAGGTAAAAATCAATTTACTATTTTTATTTGAAACATA
GAAAGATGTAAACAATTTGTACTACCAATTAATAATTATTTCGACTATTCAGTTCGATGGTTACATTCAATATTTTACGTCTAAAACAAGCTAAAATGCAATTGCGGCATTC
TGCGAAACGAGCTGAATACAAAGTAACGATTCAATTACTTTATTCAAATGTCCATTGCTCATTGATTGACCATTTTCCGACTAAACAGTATTTTCAAAGCAAACTGCGTTCG
CGTTTTGCATTCGACGACACCAGGGAATCAATTGAAGATAATGGCAATTCAATTACTCGGTATCACATACAAATTCCAAAGAAAGCCCAGAAAACTGTTCACATAGCTTTTA

GCACTCTATTTTGGGTACTTTATAGAATACAAATCCCATTAACACATAAACCGAAAATCACATTTAAATTTGATGTAGGTAATATTCCATTACCACAACTTTTTTCGACAAG
ATGCGCGTAGGCCATGGCAATTTTCTAATGCCAAACATGAGCGCTCGTTGCCCGGAAATTAAATTGATGCCGGATTATGCTGAAATTGGTGTCCCTGCGCACTGTGGGTTTC
AAAAACTGGGCTGGCTACCTTTCTATGGTGAAAATAAGCGAAAACTAATGCTAATGTTGTTGGCTTGTGTGATTTATGCGTTAAAGTCGATATCCATTTACTAGCGCAACAG
TGCGTATGGGCAATTTTGCATGTATGCATTTCATATCCGCGAAAAAAAAAATCTCGGTAATTACCCGCATTATAATTAAGCTTAACTGTTGCAACTCACCTGAAATGGTATG
CAGAAGAGGCCAATAATCACATCGGCAAAAGCCAAATTAGCGATATACATATTTGTCACGGTCCGCATTTGCCTGGTCGTGGCCACCACCCAGATGACCAAAGTGTTGCCAA
TGACGGCCACGATACTGATTCCCCCATAGAAAATGGACAACAGGGCCACAATCTCAGCGGGAGCCGCGTATAGACGCTCTGTGAATGAATTGAAAAATAGTTGTGCATTAGA
TTTATATCACTCGGAAAAAGTGCTTGCAACATTATTATAAAACTAAATGAGATATAACAATCTGTTTTACACAAAAATTTAAGAATGTATCCAATACATAATGCGCTCCTAC
TAAGTTACTGAAATTTTAGTTATTTATTACAAGACCATGCTTTTATGAGTGAATAAATTTTGTGTGTGGTGCTTAGCTTCTTCCAATCCATTAGTTTCAGTGTTGAACTTGT
GAAAAATCGCGCTTGCTGACAGAAGCAGATTAATTCAAGTGTCAGTTGTCAGTCAACTGATGGCAGGACAGACGACATTCACGTTCCTGCTGACTTGTAAACGTCAAGGAAA
TTGAACCGTATTTATAGCCAAAGAGGGAATTTTATAAATGCCACTTTGGACAGTATATACCTGTATATATATATACCTATATCCAATTTAATATAAAATATTGATAGTTAA
ACCCAACTTTGATATGAATTTCATTAGCATATCCGAACATCACCTGTTTCCTCATATCAAAGTCAATGAGATGGTTGAGCCCAAGTTTGGTCTTTAAATTCGGCACCCCCCA
CACTCATGAATGATTCAAGATCAAAGAACTAACGGCGAACTGTCAGAAGAAAAAAACACTTGTCAATGTTTTTGATTGATGATACGTCAAGTCAATTTAATCCAGTTGGATG
AGTCGCGTCCCCAGTTGCTTGAAGGGATTAGAGAGTATACTTACCAAATTCTGCTTCCTCCTCGGCTCCGGGCAGGAATTCCAGGCGGGACTCCTGCTCGATTAAGTCCATT
GCCATAACACTTTTTCGTTTTTAGTTTTAGTTGTGTATCCTAATGAGGATATGATATCCTTCAAACCTTGTATTTTAGTTCGAATGCTGCTGCTCTATGTAGTGCAATACA
TCTTCAGGGGCTTCTGAGCTTCATTGCAATCTAAACTGGAAAGATATAAATTACAGCATTTGTTCATTTGTTTTCTGAAAATTCGCAAAAAAGTTATCTGCCACAAGTCCAG
TGGTCAGGGAACAACAAAAGTGCAGACAACAAGAGAAACAAATGATCTGAAAAAAAATCATTGAAAAGCTATACAAGCAATATCAGTCTTGCCCTCAAGGAACTGACCAAAT
TCAGTGGGCGGACATGAGTTATGGCTCAGGTGGAATCACCTGTCAGCTTTCACAAGTGAAAAGTTTGTTCTGTATAATTGAAAGTGCACCCAAATTCAGTTTGGCGACTGAG
GGCACTGCACTAGAATTTCGCTTGTATTTCGTTTCTTATGAAAACAGTTCATGTTTTCATGTTTTTATTGAAGAGTCATGCCAACGCTCAAATGTTGCCTTAGCCTTATTGA
GGTGGCTCCTAAAAAATAAACAAGAAAGTCAGCAAGTTCATTGAATGTTTGAGAATAGAAGCCTCACTCAGGTTAATTGGCCATTTATCGAACTGTTAACATTGTTGGGACT
TGTTAATTTATCTTCTTCAGTGTAAACACTGTCAAGACTTGTTAATTTATTGGGTAACTAAATGTTTATTCTATAGGCAGACGGAATTGTTATAGATTGACTTAGGGATGCT
TGTAGGAAATAGGCTGGGAAATTGGCGAATGTCGTGAAAAAAAGATGTATTTAGTTTTGAATGACTGTACTATTTCCCCCAGAAAGAAACAGGTTCACAGGCTTTAAGTTTA
TATGGAAATGCAAAAATTTGTTTCGAGCATTAAATTACTCATTTTATTATGCTGGTCACAAGCTAAAAGAACTAAAAGGAGCCACCCTAACGCATGACCCCAA
(SEQ ID NO: 118)

Exon: 6383..6317
Exon: 5566..5364
Exon: 3916..3768
Exon: 3678..3557
Exon: 3487..3261
Exon: 3188..2997
Exon: 1663..1001
Start ATG: 6383 (Reverse strand: CAT)

Transcript No. : CT29768
ATGGACTTAATCGAGCAGGAGTCCCGCCTGGAATTCCTGCCCGGAGCCGAGGAGGAAGCAGAATTTGAGCGTCTATACGCGGCTCCCGCTGAGATTGTGGCCCTGTTGTCCA
TTTTCTATGGGGGAATCAGTATCGTGGCCGTCATTGGCAACACTTTGGTCATCTGGGTGGTGGCCACGACCAGGCAAATGCGGACCGTGACAAATATGTATATCGCTAATTT
GGCTTTTGCCGATGTGATTATTGGCCTCTTCTGCATACCATTTCAGTTCCAGGCTGCCCTGCTGCAGAGTTGGAACCTGCCGTGGTTCATGTGCAGCTTCTGCCCCTTCGTC
CAGGCCCTGAGTGTAAATGTCTCGGTATTCACGCTGACCGCCATTGCAATCGATCGGCATAGGGCCATCATTAATCCACTTAGGGCACGTCCCACCAAGTTCGTATCGAAGT
TCATAATTGGTGGAATTTGGATGCTGGCCCTGCTATTTGCGGTGCCCTTTGCCATTGCCTTTCGTGTGGAGGAGTTGACCGAAAGATTTCGCGAGAACAATGAGACCTACAA
TGTGACGCGGCCATTCTGCATGAACAAGAACCTATCCGATGATCAATTGCAATCCTTTCGCTACACCCTGGTTTTTGTGCAGTATCTGGTTCCATTCTGTGTCATCAGCTTT
GTCTACATCCAGATGGCGGTACGATTGTGGGGCACACGTGCTCCTGGTAACGCACAGGATTCACGGGACATAACGCTGTTGAAAAACAAGAAGAAGGTCATCAAAATGCTGA
TTATCGTGGTCATTATCTTTGGACTCTGCTGGCTGCCACTGCAGCTCTATAATATTCTGTATGTCACGATACCGGAAATCAACGACTACCACTTCATTAGCATCGTCTGGTT
TTGCTGCGATTGGCTGGCCATGAGCAATAGCTGCTACAATCCCTTTATTTATGGCATCTACAATGAAAAATTTAAGCGGGAATTCAACAAGCGATTTGCGGCCTGTTTCTGC
AAGTTCAAGACGAGCATGGACGCCCACGAAAGGACCTTTTCGATGCACACCCGCGCCAGCTCCATAAGGTCAACCTACGCCAACTCCTCGATGCGAATCCGGAGTAATCTCT
TTGGTCCGGCGCGTGGTGGTGTCAACAATGGGAAGCCGGGCTTGCATATGCCGCGGGTGCATGGATCCGGTGCTAACAGCGGCATTTACAACGGAAGTAGTGGGCAGAACAA
CAATGTCAATGGCCAACATCATCAGCATCAAAGCGTGGTTACCTTTGCGGCCACTCCGGGTGTTTCGGCACCAGGTGTTGGCGTTGCAATGCCGCCGTGGCGGCGAAACAAC
TTCAAACCTCTGCATCCGAACGTAATCGAATGCGAGGACGACGTGGCACTCATGGAGCTGCCATCAACCACGCCCCCCAGCGAGGAGTTGGCATCCGGGGCCGGAGTCCAGT
TGGCCCTGCTAAGCAGGGAGAGCTCCAGCTGCATTTGCGAACAGGAATTTGGCAGCCAAACCGAATGCGATGGCACCTGCATACTCAGCGAGGTGTCGCGAGTCCACCTGCC
CGGCTCGCAGGCGAAGGACAAGGATGCGGGCAAGTCCTTGTGGCAACCACTTTAA
(SEQ ID NO: 119)

Start ATG: 1 (Reverse strand: CAT)

MDLIEQESRLEFLPGAEEEAEFERLYAAPAEIVALLSIFYGGISIVAVIGNTLVIWVVATTRQMRTVTNMYIANLAFADVIIGLFCIPFQFQAALLQSWNLPWFMCSFCPFV
QALSVNVSVFTLTAIAIDRHRAIINPLRARPTKFVSKFIIGGIWMLALLFAVPFAIAFRVEELTERFRENNETYNVTRPFCMNKNLSDDQLQSFRYTLVFVQYLVPFCVISF
VYIQMAVRLWGTRAPGNAQDSRDITLLKNKKKVIKMLIIVVIIFGLCWLPLQLYNILYVTIPEINDYHFISIVWFCCDWLAMSNSCYNPFIYGIYNEKFKREFNKRFAACFC
KFKTSMDAHERTFSMHTRASSIRSTYANSSMRIRSNLFGPARGGVNNGKPGLHMPRVHGSGANSGIYNGSSGQNNNVNGQHHQHQSVVTFAATPGVSAPGVGVAMPPWRRNN
FKPLHPNVIECEDDVALMELPSTTPPSEELASGAGVQLALLSRESSSCICEQEFGSQTECDGTCILSEVSRVHLPGSQAKDKDAGKSLWQPL*
(SEQ ID NO: 120)

Name: Tachykinin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384568
AGTGGAAATTTCGAAATTGAAATGGACCCTTTCCGCCGAAAAATCGATTGATTTTCAATTTCCCCTCTGCGCTTTCTTGTTTTGCTAAGTTTGAAAGTTATTTGCTTTAACT
CATGCTGAAAGAAAATATCAGTTGACTAGCCGCTTTTTAACCTTTTATTCGAATTTTATCCTTAGTTGTTGCACAAAAGAGTCTATGTGCTTAGTTGATGATATCAGTCGAT
ATCTCTTTAATTGGCCAATAACTGCTGAATATGTGGCAGCACAATATAGTTACGAAAGTAATCGGCTTAAGAGCCCACACTGCGATGTGCATTTTAGTGCATTAAAAGGACT
TGCAATGCACTACTGTCTAATCCAGATTACGACGACAGCCAATGGCTTTAAGCGCAAGTATTTTAAAACTGAATTGTGAGCCGTGATTCCAGCTGGAGTTAAATAGGAAAAC
TGTTGGGCTTAAAGCAGCTGTAGTGGCAAATACCAATCGAGCATGGCTTCAAATCACGTTCATTTCGTTCCACATGACATGACATCATGGCAATGGTATCAAAGTGCACACT
TGTAGACACAGAAGCCTCGGGTTTGCTTAAATTTTTATGGCCTCAACACTTACCAGGTGCACAGCTTTTCTTGCCAGCTACTGGATGTTCAGTAGCAGGGAAATCGACAGCT
GGAAGCTCGAGTATCAAGAAAATTGAGAATGAATGGAAAATGGGTTACCACTCTCTTTGTTGCAGACAAATTTCTCCTTTTAATTGAAGGCAAAAAGGGATATATATTTTTA
AATAGTTGGCTTCTGACAGAGACTTACATGCGAAAATAAAAAATTTAGTATTAATTAAATCGACTCGCTGGGAAGTATTGTCTTATTTTATATTAGAATTTTTAAATTTTAA
TTTAGGGACCTATACATAATGGCAATCAATTAGCATAAAATATAACGTTATTAATACATTGGTCCTGGACCTTTGTTCGTCGGGGAATACGTATTTTAGTATATCTACTGAT
GTGATCTGCTGTTGTAGTTGGTGGATGAGGTGGTCATCGACTGGAGCTCAGTGGTGCGCTTTTGGACATCCAGCTCCATCTTGGTGGGGCAAATCAGACCCACCCACAGCCG
ACGGTTCGTGGAGTCCAAGAGGACAAAGTAAATGAACATCACAAATCCCTGCATAGTGGCCGTGATGCAGAAGAGGTATGAAAAAGCCACACCTGCCTGCATGAAGGCGAAT

ATTCCAAAGATCCAGGTAAGACCCAGCAAAAAGAAGAGCATTATGGAGAGTCGGATCTGCTTAACCACCATTTTCTTTTCGTTTTTATGGATGCTCTGGCTCAAAGAGTGCG
AGATGCTGTAGAAAACATACACGAATATCACCAAGTTGCACACAGTAATCAAGGTCACAGGGAGCACCACCCCAAAAATGAGACCATAGCCCGATGGATAGCAGATGCCAGT
GTCCGTGGAAAGTTGGGCAGCCGAGGGGACGTACGAATCGGGATCTATAAGGGCCACCAGCAGGGTGGGGACCAAGGGCAGTAGCCAGGCCACAATGGCGGCCTTCAGGATA
TAGCGAGGTGGTCTCTCGATTCCGATAACTGTGACGTAACGCTGGAACTGCAGGAAGGCGATGATTAGCATCCAACTGAAGAGAACCAGAATGGAGTACTGCATGGCAGCGC
CCAGAGCCACACAGCGAACTGTGTTTCCATTCACAACAAGCGCTTCGGATACATCATCTGTATTGAGAAACACGAAGAGCATCATCTGGAGACACATAGCCAGGCAAAGGTG
CAGTAAGACCTTCGTGGAGGCCTGACTGCGCCAGCTTTTGAAGAGTGCGGCAGTGAGGAATATTCCCAGTATACCCAAAAGGGAAAGACTGCAGCCCACTATCGAGATAATG
TCCAGCACTTTCTCATTAATCGGAGTAATTAGAATCTCCTCGCCAAGGTCGTTGGCTCGATAGCTTCCACCCACCAGGAATGCAAACTGTGTGAGGTGGTTAGTGTGGCACT
CGATGATGGCATCCTTCAAGAGGTCCGAACTGCTTTCAGTGCTTACTCCCTCAGTTGACCAGGTTTCATAGTTCCAATATCCGCATCCACTCCCAATGCTAAAGGCCTTTGA
ATCCGGATTGCGAAGATTTTCATTACGCAGAAGAAAGGGCAGTGGCAGAGAGAGATAATTGCCTATTTAATGGAAGGGTATTCAAATTAGTTCGAGTATTCGAAAACAAATT
TTCATGCGTTTCTTACTTCTCAGTCCTGGTATTGAAATGGAGATTACTTTGCTGCGCGGCTTACGAGTGCGTTGCAATGAGGTTTCCACGAACAGAGCATCGTGGGCATAGA
CTTTAAAGATGAGATAGGTTGCTCCCCTGCTTTTCACTTGCCGCCAGAGATTCTCCGGCAAAAATGCAGCTGTCTCCAGATTAGATTCCTCCTTGACCTTAGCCAAGTCTTC
GGAGAATCGAATGAAGCGGTACCAGAAGCCACTGGCGGAAGTCTTTCGTTGATCCCCTGGCGCAGAGAATATAGCTAATCCAGTTATGCGATCGCATTCGGGGTTGGCGAAG
AACACGCTGATATTGCCGGTAATCAGAGCTTGGACACCGATATCTGAGAAGTTGTACGTTTCCACTCCAGTAGTGGCCGTTTCTGCTTCATCACTGGTCGGCTTGGCCACCA
CCTTGCCGCATCTGTCCTTGGGAACCAGTTGCTCCGGAAGAGCATCCATGTAGCTTTCGAACTGGCTGAGCAGGGAATTGGTAGCGTTGAGCTGAGCTGATAATCTAAGAAC
CTTACTGTCGCAGGACATTATCTCCTTGCAGACACTGACCAAGTCCACACTGACCGTGGCATCCTTGTCCTGCTGCATCAGGGCGCCGAACATTTGTCCGGTCATGTGAACA
TCGGCGGGTAGAAGATTCCCACCTCTCTGTCGGAACATGTTCCTCATTATGCCAGTCATCTCACGACGCCCCTGGGTGTCATTGGTGCGCCTCCTGCCCTCGAGAACGTCGT
CATGCAGGCTATTCAAGTCACCGCTGATCGTGTGCTCCCTGAATCGCCTCAAACACACCACGGGATCCCAGTGTTCGGTGGACTCCCAGTTGGCTCGGTTGTCCTTCACTCT
GCACTCCCTGGTGACCGGCATTCCGTTTCGCAGAAGGCAAACATCGTGCAGGGTGGCACGTTCGCCGATTCTAGCCCGTTTCCACTTGTTCTGGTGAGGTGTTTGATCGTGG
TTACGCGTTATGTAGTCGTGCTGAAACTCCTCCGGCTCGCAGAACATGGACGTACAGTAAACCAGATGTGTTTGCATAAGTTCGTCCGATGAATTACGTTCCTCTCCATCAC
AGTACTGAAACGTTCCGTTTCTCAAGCACATTTGATTTTCCTTTGGGCAGGCGGAATAAATCGGCGAAACCAAGGCCAAGACCACAAAGAGCCCAGCACACCAGTTGAAATT
CATGGCGATGAATCAGGTGCCACAGCTGTTTGCAGCTCTCCGAAAATGTTTACAAATTCTTTGGCAGATAGTACTTTTCACGTCCGACGTTCTTGAAATATCGCGGCACACT
GAGCTGTCTGTCGATCGATTAGATTTCTGTAGACCCCTTCTTCATGGTGTTTGCTATCAGTTGCCACCACGAGCGCTAATCGACCCCCGTTTTTTGTATATCTACACACTAC
TCACAAAAAGACCTCAAGTCTGCGGCATTTGATACAATTTACATTGTGATTAACACGACCACGGCAATTGAAGGCGAATCCGACGGAACCCTTTGAGGCAGGGAAGTAACAC
CAAACTTGTGACGTCGTCTTTATTTGTGCAAATAGCGATTTCACGCCTTTTTACCGGTCTTGGTACTTAAAATGCTCTGGCTGATTCATTTTTTACGCGCTTGAATTGCCAG
ATAAGAAGTTCTTACCAGGAAACACGGCTGTACATGGGCTATTTATACATCAATGATACAAAAATAATGTTAAAAAAAAATTTAAAAACTTTAAAAACTAGTTTTATTTGCAT
CCGTTCAAGAAAGAGATTTTTTTTAGATTTGTTTGATCAATTCATTGCTATGGCTTCTAATAAATGTGAAATGTGAAACACCTTCACATAAAAAAGTATTTACGCAATTCAC
TCAAGCAAAACAGATCCTTTGTCCTTGAAAAAAGTTATGAAATGCAAAGCACTTCACCTTGGTTGCATAAAGTGCGATTTTCATTTTGACAGGCTACGCGTTTTCAAAGACT
CAGCCTTCGATTATAATCCACCTGTACTCACTTCATTGTAAGCCAGTTTTTGGGCCAAAAAGTAATTTTTACAGCCGACTTCGCAGCACGTAGACTGCGTGTGCTAGCGGAA
ATTGTTAGCTTGTTGACACGTTT
(SEQ ID NO: 121)

Exon: 3391..2257
Exon: 2190..1001
Start ATG: 3391 (Reverse strand: CAT)

Transcript No. : CT31591
ATGTGCTTGAGAAACGGAACGTTTCAGTACTGTGATGGAGAGGAACGTAATTCATCGGACGAACTTATGCAAACACATCTGGTTTACTGTACGTCCATGTTCTGCGAGCCGG
AGGAGTTTCAGCACGACTACATAACGCGTAACCACGATCAAACACCTCACCAGAACAAGTGGAAACGGGCTAGAATCGGCGAACGTGCCACCCTGCACGATGTTTGCCTTCT
GCGAAACGGAATGCCGGTCACCAGGGAGTGCAGAGTGAAGGACAACCGAGCCAACTGGGAGTCCACCGAACACTGGGATCCCGTGGTGTGTTTGAGGCGATTCAGGGAGCAC
ACGATCAGCGGTGACTTGAATAGCCTGCATGACGACGTTCTCGAGGGCAGGAGGCGCACCAATGACACCCAGGGGCGTCGTGAGATGACTGGCATAATGAGGAACATGTTCC
GACAGAGAGGTGGGAATCTTCTACCCGCCGATGTTCACATGACCGGACAAATGTTCGGCGCCCTGATGCAGCAGGACAAGGATGCCACGGTCAGTGTGGACTTGGTCAGTGT
CTGCAAGGAGATAATGTCCTGCGACAGTAAGGTTCTTAGATTATCAGCTCAGCTCAACGCTACCAATTCCCTGCTCAGCCAGTTCGAAAGCTACATGGATGCTCTTCCGGAG
CAACTGGTTCCCAAGGACAGATGCGGCAAGGTGGTGGCCAAGCCGACCAGTGATGAAGCAGAAACGGCCACTACTGGAGTGGAAACGTACAACTTCTCAGATATCGGTGTCC
AAGCTCTGATTACCGGCAATATCAGCGTGTTCTTCGCCAACCCCGAATGCGATCGCATAACTGGATTAGCTATATTCTCTGCGCCAGGGGATCAACGAAAGACTTCCGCCAG
TGGCTTCTGGTACCGCTTCATTCGATTCTCCGAAGACTTGGCTAAGGTCAAGGAGGAATCTAATCTGGAGACAGCTGCATTTTTGCCGGAGAATCTCTGGCGGCAAGTGAAA
AGCAGGGGAGCAACCTATCTCATCTTTAAAGTCTATGCCCACGATGCTCTGTTCGTGGAAACCTCATTGCAACGCACTCGTAAGCCGCGCAGCAAAGTAATCTCCATTTCAA
TACCAGGACTGAGAAGCAATTATCTCTCTCTGCCACTGCCCTTTCTTCTGCGTAATGAAAATCTTCGCAATCCGGATTCAAAGGCCTTTAGCATTGGGAGTGGATGCGGATA
TTGGAACTATGAAACCTGGTCAACTGAGGGAGTAAGCACTGAAAGCAGTTCGGACCTCTTGAAGGATGCCATCATCGAGTGCCACACTAACCACCTCACACAGTTTGCATTC
CTGGTGGGTGGAAGCTATCGAGCCAACGACCTTGGCGAGGAGATTCTAATTACTCCGATTAATGAGAAAGTGCTGGACATTATCTCGATAGTGGGCTGCAGTCTTTCCCTTT
TGGGTATACTGGGAATATTCCTCACTGCCGCACTCTTCAAAAGCTGGCGCAGTCAGGCCTCCACGAAGGTCTTACTGCACCTTTGCCTGGCTATGTGTCTCCAGATGATGCT
CTTCGTGTTTCTCAATACAGATGATGTATCCGAAGCGCTTGTTGTGAATGGAAACACAGTTCGCTGTGTGGCTCTGGGCGCTGCCATGCAGTACTCCATTCTGGTTCTCTTC
AGTTGGATGCTAATCATCGCCTTCCTGCAGTTCCAGCGTTACGTCACAGTTATCGGAATCGAGAGACCACCTCGCTATATCCTGAAGGCCGCCATTGTGGCCTGGCTACTGC
CCTTGGTCCCCACCCTGCTGGTGGCCCTTATAGATCCCGATTCGTACGTCCCCTCGGCTGCCCAACTTTCCACGGACACTGGCATCTGCTATCCATCGGGCTATGGTCTCAT
TTTTGGGGTGGTGCTCCCTGTGACCTTGATTACTGTGTGCAACTTGGTGATATTCGTGTATGTTTTCTACAGCATCTCGCACTCTTTGAGCCAGAGCATCCATAAAAACGAA
AAGAAAATGGTGGTTAAGCAGATCCGACTCTCCATAATGCTCTTCTTTTTGCTGGGTCTTACCTGGATCTTTGGAATATTCGCCTTCATGCAGGCAGGTGTGGCTTTTTCAT
ACCTCTTCTGCATCACGGCCACTATGCAGGGATTTGTGATGTTCATTTACTTTGTCCTCTTGGACTCCACGAACCGTCGGCTGTGGGTGGGTCTGATTTGCCCCACCAAGAT
GGAGCTGGATGTCCAAAAGCGCACCACTGAGCTCCAGTCGATGACCACCTCATCCACCAACTACAACAGCAGATCACATCAGTAG
(SEQ ID NO: 122)

Start ATG: 1 (Reverse strand: CAT)

MCLRNGTFQYCDGEERNSSDELMQTHLVYCTSMFCEPEEFQHDYITRNHDQTPHQNKWKRARIGERATLHDVCLLRNGMPVTRECRVKDNRANWESTEHWDPVVCLRRFREH
TISGDLNSLHDDVLEGRRRTNDTQGRREMTGIMRNMFRQRGGNLLPADVHMTGQMFGALMQQDKDATVSVDLVSVCKEIMSCDSKVLRLSAQLNATNSLLSQFESYMDALPE
QLVPKDRCGKVVAKPTSDEAETATTGVETYNFSDIGVQALITGNISVFFANPECDRITGLAIFSAPGDQRKTSASGFWYRFIRFSEDLAKVKEESNLETAAFLPENLWRQVK
SRGATYLIFKVYAHDALFVETSLQRTRKPRSKVISISIPGLRSNYLSLPLPFLLRNENLRNPDSKAFSIGSGCGYWNYETWSTEGVSTESSSDLLKDAIIECHTNHLTQFAF
LVGGSYRANDLGEEILITPINEKVLDIISIVGCSLSLLGILGIFLTAALFKSWRSQASTKVLLHLCLAMCLQMMLFVFLNTDDVSEALVVNGNTVRCVALGAAMQYSILVLF
SWMLIIAFLQFQRYVTVIGIERPPRYILKAAIVAWLLPLVPTLLVALIDPDSYVPSAAQLSTDTGICYPSGYGLIFGVVLPVTLITVCNLVIFVYVFYSISHSLSQSIHKNE
KKMVVKQIRLSIMLFFLLGLTWIFGIFAFMQAGVAFSYLFCITATMQGFVMFIYFVLLDSTNRRLWVGLICPTKMELDVQKRTTELQSMTTSSTNYNSRSHQ*
(SEQ ID NO: 123)

Name: HE6 Receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384531

GTGTTAATCCCCTTGGCTTATCGGTTTAATCTATTTATAAAATCAATTCAAACATAAAGCGGCGCATATCTCAGCGATGTTGGGTAGGTAACCAAAAGTAAATACACTCTAA
TTCCTATTGATCTTGTATAAATACGAGTGGATAGTTAGTTAGATATCACCATTTTGCACCGTGCATAAACATAAAGTGAGATACTCGTTTTCGAGAGATGCACTTCTGAGTA
CCCGTATTCAAATCGTACTCTCAGTAGCCAAGTACTCAGTACTCAAGCTGTTTGTCTTACGGGTGTGGACGTGCGAGCAGACACTTTAGGAGTGGGATAGCCGGAGCAGACC
GGCTGAAGTGTCATCCGATCGAAGGTCGCCTCAACAGCACAGCGATTTCGATTCCGATTCCAAATCCGATTCTTTGGCCATAGGAATTCCTCTGGGAGGTGGAGCATGTGGC
TGATTTATTTATTTCGAGCGATATAAACAGAGCAATACGAGAATGGGCGAATATAAGTTTGCTCGAACGGCTGGTCCGTTCCGAGCTGAAATTGAATTGTAAATGGCCATAT
TAATTGGGAATCATTTAAAATTTCCAATGATTTTTTCTGTGTTTTGGTGTCGGCAGTCTATGCATGTTCGGTGGCATAATTTATGGGCGCTAATTACTATCGCCCCATTTTG
TGTATTTCGCGTGCTGTCCAAGGCCAATTACGCTTACTTATTTTGCAATTTATGTGGCCACAGTCCGGCCGTAATTGGAAATTGTGTAAGGTCGCAGGCAATTTCATTATTTGT
ATGCACTTCTGGACGCTCCGGTTGCCGGGTCTAGCATTTTTCTAGCATTTTTCTAATGACGCCAGCGCGTTTGGGGACGACGTTACCTAGCCCAGTTTTCCCTTTGCAGCCA
TTTGATTAAGTGACCTACGTTTCGCCCACACAGAATCATCAATTTGTTCATGTCGCCGACGGTGCGAATAAGGTCCAACCTAAGTGAGTGAGCTATCCGCCGGCATGGCGGC
GTTGGAACATCCGGATGCCATGCTGCCATACTCCAAATTGTTGTCAACAGCCCTCACTGTGGCCAGTGGTAACTCGAACTCCTCCGATGGTACTTTGCTGCTCATCCAGACC
AGTGCGGCACCAGGTGGCGATGTGATGGCCACCGGAGCCGGAGCCGGAAAGTTGGGTGTGCCCTCGATGAACTCGTACCTGGTGAGCATGGCGTGGCCCAAGTCGCTGGCAG
TGGCCGTCTTCCTGGTCATCATCCTGGTCACCGTGGTGGGCAACACACTCGTCATCCTCGCCGTTCTGACCACACGCCGCCTGCGCACCGTAACCAACTGTTTCGTGATGAA
CTTGGCCATCACCGACTGGCTGGTCGGCACCTGCGTGATGCCTCCATCGGTGGTGCTCTACATAACAGGTGAGAATGGCAGCGGAGGGCCTGTCCATGCCCATCGATGGCCA
AAGATTTATGGCTCTGTGGGGGCTTTACACTTCCGGGGAAGGTGGAATCGGGGCAAGTGTTTACTATTCGAAGCCCTCCGGTGGAAGTTCGGGTAGTAAACATTTGCCCAAA
TCGATGGGGGAGCACATTTGGTCGATTTCGATTTAAAATAATCCCACAGCGAAAATTATAAAATTTCTCCGTTGGCAAAATATACATATTTTTGATTGCAAAATATCTTGTG
AGTGGTGTGCGAGTGTTTTATGGGTTTCGCTGGAAAACCTTTGATAATTTTTCTGGAATCTATATTTTAGCAGAATTTTGAATAAAGCATAACAACATTAGTGGTTAAAA
ATTTGTATAATTTAGTAATTTTAAATAAATTATGTATCTTTAATAATTAACATCTCTTATTTAACGCATTTCATAAATGTATAAAAAGCAATCAAACATGTGATCTTTTCG
AAGCCAAATGAATATACAAAGTTTAATATTCCTTGTCATAACTTGAACAATACAAAGAAGAATTTATAGTGGTTTAAATTGAAGGATCTGCCAGTGTAATGCCTCCTTAATC
CCGGCCAGGTGCACAAACCGATTGAAGCACGGCGACAAATATTTGTTCTTTGCAGGCACTTGGCGCTTTGGCTGGATACTGTGCGACATTTGGATCTCGCTGGACATCCTGC
TCTGCAGCGGCTCCATCCTCAGCCTGTGCGCCATCAGTCTGGACAGGTGAGTTGGCCAACTAATTTCGTAATTTGCCGTTGATGTCATAGCAATGACCACAGTCCATAGCCC
GTTGTCCACGGTCCTCGGTCCTTAGATAAAGTCATGGCAAGGTGCAAAAACAGTGCAAATAATTTAATTGTTGACGTTTTTATTGCCGCATGACCGTTGTCCTGGCCATTCC
CAGTGCATATCCTTTGAGTGTGCGACGCCACAATGGCAAGGATTGGGTTGGGCTTGGATTCCCATTCACCCATTCCCTATTTCCCCATTTCCTCCATTTAGCCATCCCACCG
TTTGCTTTTGTTGGAAGTTTCAATTACATTTTAATGGCATTTCCAAGTTTTGTGTTGTGGCTCCCGAGTTGAGCTGATCCAGCTGAGCCATGGAGCCATGGACCTGAACCTC
AACCAGGACAAACACCGACGCCCACTGGCGTTGTGTGTGCCTTTTGGCCTGGCCAGAAAGCCGTCTGCGGCTTTGCGAACAATATGTCCGCTTGGCTTTTGTTTTTTTCCAG
CATTTCTCTGGCCATTTCCGTTGCGTTTATTGGCCCGGCTTTTATAATTTCCTCTTTTCACCTTCGTTTTTGTTTTATATTTCGCAGGTATCTCGCCGTCACACAACCACT
GACGTACTCCAAGAAGCGCCGATCCAAGCGACTGGCCCTCCTCATGATTCTCGTCGTGTGGATAACTGCCCTGTCAATCACATGTCCACCCTATTTGGGCTGGTAAGTTAGC
GCCAGCTGCACGGAGAACAAAATCAACCTTAACCCATGATGGATCGCCTATTCATCTGTTAAGAAAACCATTCATGTTATAATCATTTAATTGAACGCATTAAATTGGGTTG
GGAGTACAGGCAACATGTGATAATGATCAGTATCCAAAGGAAGAAAAAATAATATATTTAATATCATATAATATTATTAATAATATTATATTATAAATATGGCCAAGATAAAT
TTATTTGTAGTGCTCTATTAAGTTTATAGATATATATTTTTTTGAATTTTGTAATATTATTTTCACTCTTTCCCGTGTAATTCGAAATATTATATGTAATATTTTTATTGAG
CACCCATTTTTTTTTACCAAATTTGTCCTTCCGGAGGCGTGACAATACTTTGTGGCTTTTTGTGGCTGAAAGTCCTGCTCGCTTTTTTTTTTCTTTTTTGAGCTAATGGCT
GTGCGGTGGCGCAGCATAGATAATGTGCTACAAAAGGGTCTCGAAATACACACGTTTATTTATATTATATTTGCGCCTTTGGCTGCGCCTAAAAAACATTACGTAATTGACG
ACGGGTTGTTGCTGTTGCTGCTGGTCGTGCAGCCATTACATTGAAGTTTTAATGAGTTGTTAATGACTTTGTTTGCCGCGGCGCATTTGCTTTGTTCTCACCTACGCACACA
CACTCAGCGCGAAAGCCGCACACTTGTGTGTTGCATACTTCCAGGCGCCCAAATTAACCAGCACTCACTCACACACACTCGCACTCACTCACACACACCACAAAGCCAGCAT
GTCCTTAATGCCATTGGCCATGGCTTAGTTTTGGCCCCCTTTGGCTGAAAGCGAAGAGCATGGCTTAAAGGGCACATTAAAACGGTCTGAGGCCGAATTACAGCCACAGCTG
CAGGACATTGGAAAATTTGGAAAATTTGGCGCTGGGGAAAATCGCAATTATAAGCGCCAGCAGATGGCGCCACTTACATATTTTCATTCTTTAAGCGGACTAGTTTTATGA
AAAAAGTTTGCAAATGCAGCGGTTGCAACAACTGTTTGTTTTACTCACAATCTGTTATTTGTTATTGCCTAATGAATACTAAACTGATTGTGTTTGCTTGCTATTTAAATTT
AGATTTAACTTTTGGCTAATTGATTTAAGCATTGTTTATTTAGTTCATTTTTTAATTTAACTAATGCATAGCTAATGTATTAAATTGCCCTCAATTTCGGATTGAATAATTA
ATGTTAAACAAGAAGTTTAGTAGGAAAATCGGCTATTTCCATTCGGGTCTTGTAAATACCAACCTTAACAAAGTTGCATTTGTACAATTATACTGTCAGGTTATTTTGGTAT
CGGGTAGCAGTAATTACCGTGAGTTTTTGGATAACACATACTTGGGGTAACAAGCAGCTAGTTTAGTTGAATTACAACACCGGAAAATTCCATTGACAGTAGTGACAATGAG
GCAATAAACTTTTTGAAAGAGAAACATTATAGGAAACTGTGTATCCAGTGGGCGTGCCCCAAACTTTGGTTGGCATATTTGTAAAAATTGGCCATCCAAAAAATAAATCAAA
ACAACTGGCAATCATTTTATAATCGTGTGGGCGTGAGTCTTTTAGGCGATTGGGGGGCTTTAGAGTGGGCGTGGTAATGTTCTGAAATAAGCTTGTTCCCGAGATTTGGACT
CCGGGGGACAAATGGACATAACCGAATCAAGGATATATATACTTCATGCTTCTGTCTGCTACATGATATTCAGTGAACCTAGTATTCTAAGAATAGTGGGTACAATAACCCA
GTATGAACTGGAAGAAGCTGAATCTTGCCACCTACAGTGGATGTATGTGACAATTTGCCCATTTTTTTTACTCAATTCCCCACAGCACACGGGTGTACCTTTAGTTGACACA
TTTGGTGTACGCCACCACATTACTCTGTTGTCTCGGAATCGTTTTCTTCCGCATATTGATTTAAAGTCAGCTCCGATTTGTACTGCCTGAATGACCAACACACTGCGATGTG
CATAAGTAAAAGGCTTACAGGTGGGACTACTCTGACTGGCTTAGGTTTAGGATTTGAATTCGAATTTCGCCCGGCTTAACTCAGCCAGTTGTCATTCCCCAAGTCCCAGTTC
CCCAGCCGCAATCACAACCGCAAGAGCCACCAAACCAGAGCAGGTACCAGGTACCTCCCACATCCACCACATCTCACATCCTTCCAACCCGCAAACCGCATGCAATTCGAAC
TGGAATTCAATTTGCCATTCACATCCACATTCGCAGGTACGAGGTCGGTCGGCACCAGGCGGAGTTCGTGGACTGCCGCTACAACCAAAACAAGGGCTACGTGGTCTTCTCG
GCGATGGGATCGTTCTTCATTCCCCTCACCGTGATGCTCTACGTTTACGTGAAGATTGGTTATGTCCTCACGTCACGGCGACAGCGCATTGTGCGCGATGCGGTAAGCTACA
TCTCCAGCTCCATCACTATATAAGCATATATGTATGTGTCCACCACTTTGGTGTCCCCTGAAATGTGTCCGGAAATTTATCATGTCACACCTAACACATCTAACCAGCTAAC
CAGGGGACTAAGTTGTCAGGCACTGGGGAAAAAGATCCAGGAGTGCGCTCAACACTTCAAGTGCAGATCTTCATTAAGTTGGGATACCCACTGGGTATCCCACTTGAGTACC
AGCCACTTATCGCACAGTACTGCAGGTGATATTTTAAGTGCTGAGCTTTGATAACAGAAGTCCAACTAATATATCTAATCACAAATTAAATTTGCTGTGAATTCTTGGATGT
AAACTGAATGAATTATCTTTTGTTTGTTAGCATGAAAAATGAACTTCAGTGTCAATAAATGAAATGAAAATGCAATAAGAACACTTTTTGAATATTTAAAAAGAGTGCATAG
TAGTTGATAAATATATTCTTTTAGTTTAGTACTAATTGAACGTTTGCTGCCAAGAGCAAGTGAAGTAAATTGTAGAAAGTAGGTGACGTGAAATGGAACGAAGTAATTACAG
CAATCCAAATCCCGAATCTCTATGCCCTGACATGCAATCTAAATTATTCCTCACGTTTTTGTCAGTGCTCTGCTTACATATTCCTCTTCTATTATGTCCTTCCAACAGTACT
CGGAGCGCACGGCAGACTACGATGTGGATGGGGACAACTTCATCTCGGAGTCGGAGCACTATCATTGCACGCCAACCAAGTGGCTGCCGAACAGGAAGTCCCGCTGGAGGTT
CAACTCCCTGCACGACCCACCACCCAGCGCCGCCACGAATGCAGCAGCGAAGGGACACCAGTCGTCCGTCAAGTGCTCCAAGTGCTCCAAGTGCTCCTCTGCCACTGGAACT
GCTGGAGCTGGAGCTGGAGCTGGTGTGGTAGACAAAACGCTGACGTTCAAACATCAGCCCACCTTTTACGAGCTGGTCGAGGTGTCGCGTCTCTCGTCGCTCATTCACTGCT
CGGCAATTAGCTGCAAGTATGGAGGACCCTGCATGCACTCCACGCCCTCGGGCATGCACTACGTTGGCACGGAGGCCTCCTTCTCGGACACCTGCCTGGGTGGCAATGTGAG
CAACAACACCACCACCACCACCACCAACACCGTCGCACTGGAAACCAACGAGAAACTGGAGGCGGAGGCCAAGCAAAAGCAACAGCAGCAACGATCCTCCTTCGACCACCAT
CACCCGCATGCACACCATCATGGCTCCCATGGCCAGCAGAGCCATCACAATCATCACCATCGCATGCCGATGCGAGTTTCGACCACGAAACGTGATAGCAAGACCGCCAAGA
CCCTGACTATTGTGATGGGCGGCCTAATTGCTTGCTGGCTGCCCTTCTTCGTCTATTATCTGCTGATACCCTTCCTGCCGCGACCCGCCGTCCTCGAGGACCTCATGTTCGG
CTTCACCTGGATTGGCTGGGTCAACTGCGCCATCAACCCGTTCATCTACGCCTTCTACAATCCGGACTTTCGCACTGCCTTCTGGCGGCTCACCTGTCGCCCCATTTGCAAG
CAGAAGCGTCCGCCCAACCACCTGGCCATGTTCCGTGGCTAGTTTCCGGATGCAGATGCAGTTGCGGATGCGGATGCGGATCCGGAACCGGTTTCCGAATCACTGAGCACTT
GAGTGTGTGTCATAGTGCCGCCATCTCCAGTCGCACAATCGACGCTCACTTGGGTGGTGGGGAAATATCTCAGGCATATCCTTAACTTTTCGCCGTGCTGGTTCACCAACT
TCTATGTAATTAGCATAAAAAATAAGTGGAGAAACTACCAACTCGCATTTCGATCTTTTTGAAGTACAAGTTATGCCAATGCAAGGACATCAAATTGACTCTAGACTACAGT
ACAGATAGTAAATAAAAAAGAATTTAAGGAGTATTAACTACAATAGATAAGCTGTTTGAGAAATGCACTATCTCTTATATTAAATCAATTGAAGATAATGAAAATCAAATAT
TAGCTTAACAAAACTGCTGAATTTAATTCAAGTTTATTTCATGAAAACAAATATATAGAAACTAGTTTAAAATTTATCACACATGTCCCTCCAAAAATTGTAACTAATATTT
TGGCAAACGAAGCCAATGGAGGGACCTTTTGGCCAACTCGTCGCCACTAGCTCACCAACTAAGTATTCATTTGCCAGGCACTATAAGCTGTTTGCCGCACTCTATCTGCATC
GTTTATGCCAAACGCAGCTAGTAAAACTAAGATACTTATGGTCAATATTGATTGGAGCAAAGATTGCCTTGTTAAACCCAAAAAATCGACTGCCTTTTGAAGCTTCAAAAGC
CGCATTTCGAGCAGCATCTTAAAGGTCACTTTACGGTCGCAATTTTTTTTTTTAAGACTCACTGGACTTGGCCGGAAGTTGAGATTCTTGCCTAGACCAGATCCAAATCCT
ATATATAGTATTTTCCTGTGTGTGAGCGTTGATTGTGCGAGGCGAGCGGCACGAAGCGAAACGTAACGAAACAAAAGACAAATTATAGTATTTGCGAGTCTTTGTTGCCTGA
CGGCAACTGTTGGCCATCGGGATGAGTGGGATGA
(SEQ ID NO: 124)

Exon: 1001..1412
Exon: 2072..2174

FIGURE SHEET 57

Exon: 2777..2902
Exon: 5189..5366
Exon: 6045..6491
Exon: 6549..6986
Start ATG: 1001

Transcript No. : CT37292
ATGGCGGCGTTGGAACATCCGGATGCCATGCTGCCATACTCCAAATTGTTGTCAACAGCCCTCACTGTGGCCAGTGGTAACTCGAACTCCTCCGATGGTACTTTGCTGCTCA
TCCAGACCAGTGCGGCACCAGGTGGCGATGTGATGGCCACCGGAGCCGGAGCCGGAAAGTTGGGTGTGCCCTCGATGAACTCGTACCTGGTGAGCATGGCGTGGCCCAAGTC
GCTGGCAGTGGCCGTCTTCCTGGTCATCATCCTGGTCACCGTGGTGGGCAACACACTCGTCATCCTCGCCGTTCTGACCACACGCCGCCTGCGCACCGTAACCAACTGTTTC
GTGATGAACTTGGCCATCACCGACTGGCTGGTCGGCACCTGCGTGATGCCTCCATCGGTGGTGCTCTACATAACAGGCACTTGGCGCTTTGGCTGGATACTGTGCGACATTT
GGATCTCGCTGGACATCCTGCTCTGCAGCGGCTCCATCCTCAGCCTGTGCGCCATCAGTCTGGACAGGTATCTCGCCGTCACACAACCACTGACGTACTCCAAGAAGCGCCG
ATCCAAGCGACTGGCCCTCCTCATGATTCTCGTCGTGTGGATAACTGCCCTGTCAATCACATGTCCACCCTATTTGGGCTGGTACGAGGTCGGTCGGCACCAGGCGGAGTTC
GTGGACTGCCGCTACAACCAAAACAAGGGCTACGTGGTCTTCTCGGCGATGGGATCGTTCTTCATTCCCCTCACCGTGATGCTCTACGTTTACGTGAAGATTGGTTATGTCC
TCACGTCACGGCGACAGCGCATTGTGCGCGATGCGTACTCGGAGCGCACGGCAGACTACGATGTGGATGGGGACAACTTCATCTCGGAGTCGGAGCACTATCATTGCACGCC
AACCAAGTGGCTGCCGAACAGGAAGTCCCGCTGGAGGTTCAACTCCCTGCACGACCCACCACCCAGCGCCGCCACGAATGCAGCAGCGAAGGGACACCAGTCGTCCGTCAAG
TGCTCCAAGTGCTCCAAGTGCTCCTCTGCCACTGGAACTGCTGGAGCTGGAGCTGGAGCTGGTGTGGTAGACAAAACGCTGACGTTCAAACATCAGCCCACCTTTTACGAGC
TGGTCGAGGTGTCGCGTCTCTCGTCGCTCATTCACTGCTCGGCAATTAGCTGCAAGTATGGAGGACCCTGCATGCACTCCACGCCCTCGGGCATGCACTACGTTGGCACGGA
GGCCTCCTTCTCGGACACCTGCCTGGGTGGCAATAAACTGGAGGCGGAGGCCAAGCAAAAGCAACAGCAGCAACGATCCTCCTTCGACCACCATCACCCGCATGCACACCAT
CATGGCTCCCATGGCCAGCAGAGCCATCACAATCATCACCATCGCATGCCGATGCGAGTTTCGACCACGAAACGTGATAGCAAGACCGCCAAGACCCTGACTATTGTGATGG
GCGGCCTAATTGCTTGCTGGCTGCCCTTCTTCGTCTATTATCTGCTGATACCCTTCCTGCCGCGACCCGCCGTCCTCGAGGACCTCATGTTCGGCTTCACCTGGATTGGCTG
GGTCAACTGCGCCATCAACCCGTTCATCTACGCCTTCTACAATCCGGACTTTCGCACTGCCTTCTGGCGGCTCACCTGTCGCCCCATTTGCAAGCAGAAGCGTCCGCCCAAC
CACCTGGCCATGTTCCGTGGCTAG
(SEQ ID NO: 125)

Start ATG: 1

MAALEHPDAMLPYSKLLSTALTVASGNSNSSDGTLLLIQTSAAPGGDVMATGAGAGKLGVPSMNSYLVSMAWPKSLAVAVFLVIILVTVVGNTLVILAVLTTRRLRTVTNCF
VMNLAITDWLVGTCVMPPSVVLYITGTWRFGWILCDIWISLDILLCSGSILSLCAISLDRYLAVTQPLTYSKKRRSKRLALLMILVVWITALSITCPPYLGWYEVGRHQAEF
VDCRYNQNKGYVVFSAMGSFFIPLTVMLYVYVKIGYVLTSRRQRIVRDAYSERTADYDVDGDNFISESEHYHCTPTKWLPNRKSRWRFNSLHDPPPSAATNAAAKGHQSSVK
CSKCSKCSSATGTAGAGAGAGVVDKTLTFKHQPTFYELVEVSRLSSLIHCSAISCKYGGPCMHSTPSGMHYVGTEASFSDTCLGGNKLEAEAKQKQQQQRSSFDHHHPHAHH
HGSHGQQSHHNHHHRMPMRVSTTKRDSKTAKTLTIVMGGLIACWLPFFVYYLLIPFLPRPAVLEDLMFGFTWIGWVNCAINPFIYAFYNPDFRTAFWRLTCRPICKQKRPPN
HLAMFRG*
(SEQ ID NO: 126)

Name: Octopamine receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384349
TCCATGCATTTATAATTCAACATTCGCTCGCGGCCTTATGAAAAGCTCATTTGGTTACGAGTGTCCTGGGTCCTTAAGTCCTTTCTACCTCGAAAGTGTTAATGAGCCGGCA
GGCTGCCGTAATTATGCAAATTTTTAATGACTATGAAAATGGTTTCCGCGGAGCAACAAGTGCTCGCGTCTTGGTCCTTGGATCTCTATACCCCTCATACCCCTCTAAAAAT
GCCAGCTAATTCCATTTGACCCCAGAGGCCCCAAAGCAATCCATTTGCATGGCTACCAATAGCTCGTCGCATTCGTCGTACCAACCCAACCATTTTCCATTTAATCCCTACG
ATGTGAGAGCTCTGTACTGAGCTTACTGCTGGCATTCGGTCTTTAGATTTATACAGATTTTGTCAAACAAAACATGAATTCGCACTTTTACACACATATGTGTGTCTGTAAA
TGACTGAAAGGGAGAGCGTCGAAAGCCGATTTGGTGCCATTCAGTTGAAGCCCTTTAGGTCGTGATTCGCATATGGTCGAGGCCATTTAAGCTGATATCAGGCACATTAAAT
GTAGTAAAGGAACGTAAGTATGATAAAGGCAGCATCACAAAAGTGTCAAGTAAATGCAAATTAACATTAATTTAAGACTTGAATTCGTCTGTACTGAAACAGTACAATTAAA
ATATAATTAGCTATATATTTGTAACTTTTAAGAACGGATACTTCTGGGACCAATGTTTTTGAACATAACAGATATCTACCATATGAAGTGAAATAAATAGAATTTTTCGGG
ACCATTTTGTTTCCAGATTTCAAATGAAATGAAACGCATATTTTTGCCAGTTTTTGATAATAAATATCAGGGGACTTGGCCAAATTTTAATACAATCTCATTTAGAAAAATA
AAACAATTGCATGCCGATAAAATAATAAGTGATACCGAAAGAGAGCTTTTGATTAGAGTGAAATATGCTCATGCGGTTCAGTCGAGATCCAAAGTTTGCGCAGAATGCTTCT
GAATATTTTAGCTATAATTCTTGTATTCGTAATTTCAAGCCAATCGGAGGCAGTTATTCCTGGATGTGATTACTTCGACACCGTGGATATATCTCATATCCCAAAGCTAAAT
GACTCCTATGCGTACGAGGAGCTAATCATTCCTGCTCATCTCACGGGTCTATATACATTTAGGCAACTGGCGGATGGATCACAGGAGCCAGTGAAGAGTCACTTAAGAGCAT
GTATCTGCAAACTAAAACCTTGTATCCGATTCTGCTGCCCCCGCAACAAAATGATGCCGAATAGTCGCTGTAGCGATGGACTCACAGAGAATCTGAAGCGGATCAATCCCTA
CTTGAAGATCACACTGGAGGATGGAACTATAGGAAAATATTACCTTCTCACCGACATGATCGTCTTAAGATATGAGTTTCGGTATTGCGAGAAGGTTGTTTCGGTACAGGAG
GATCAGTACAAGCTATATGAGGTTAAGATTTCAATCTTTTTGAAGTATCTCAGTATAATTTGATATCTATAGAATGGCAGCTTTATGATTAAACCCGATGTAAACTGGACAT
TGAGCAAGCAGTGGTACTGCCTCCATCCCCGTTTGGAAGATCCGAATTCCATATGGATACTGGAACATGTATATATACCGAAGTCAATGCCAGCCGTACCTCAGGGCAAGCA
CCTACTATATCATCTAATCTCATCTCATCTAATTCTCTGATTTTTCCAGTCGGTACTATTTCCATGGTTGGTTGTATTCTCACGATTGCAGTGTATCTCTACATAAAGAAAC
TGCGAAATTTGCTTGGAAAATGTTTCATTTGCTATGTATTTTGTAAGTTTGTGCAGTACCTTATCTGGGCAGGAGGAGATTTGAACTTGTGGAATAATATTTGCTCCCTAGC
AGGTTAGTTGAATCCATATTAAAGTTTGACTTAAAATATGCGTTTATTCTAAACAGGTTACACTAACTATTTCTTTGCTTTGGCTTCTCACTTTTGGCTCTCTGTTATGAGT
CATCAAATATGGAAAAACTTGAGGTTAATCAATCGGGATGAACGTAGCTATCACTTTCTCATCTACAACATCTACGGTTGGGGCACACCTGCTATCATGACGGCAATTACAT
ATCTGGTGGATTGGGCCTGGGAGGATAGACCTGACAAACTGAATTGGATTCCCGGTGTTGGCTTATATCGATGCTGGATAAACAGTGGGTTTTAGGGTAGAGTGTTTAAGCC
AGTTTAAGACCAGTAATTGATTTTCTTTGAACAGCCTATGACTGGTCTGCTATGATATACCTATATGGACCGATGCTGATCTTGAGTTTATTCAATGTGGTCACGTTTATCC
TGACAGTAAATCACATAATGAAGATTAAGAGTAGCGTTAAGAGTTCTACCCAACAGCAGCGAAAGTGTATACAGAATAATGAGTAAGCTTTAAGTATTTATAACTTACTATT
TCTAATTCCATTCTCTTACTATAGTTTTCTATTATACCTACGACTGTCGGTGATGATGGGCGTGACTGGGATTTCTGAGGTAATAACTTACTTCGTCAAGCGGCACAAATTT
TGGCGACAAGTCCTTCGGGTGCCAAATTTTTTCCATTTGGGTTCCGGTATAGTCGTATTTGTGCTATTTATTCTTAAGCGCAGCACATTTCAAATGATAATGGAGAGGTAAG
AGCTTTTTGCTTAATAGATTTTAAAGTTACTATATCAATATCATTTTAGAATCAGCGGTCCGAGGAGACAACAACCTGCATCTTGATCAACCTACTAGTTTAAATAAATGCA
AATTATAGATAAGTGAGCTCTTTTTTTTTATAAGTTTAGTATAGAGCAGTCCCTCTATGGTCAATAAATGCTTATAATCTTCTTATTATTGAAAAATTCTTTATTATTGATA
TTTAATCAATAATATATTTTTATTTTTTTTTTTTTGTTAGAACAAAATGTCAATTGTCATCAAAAATATGTGTTTGTGAATAAAATCAAGTTCCGTAGGAAAATATGCACT
TTTATGGAAACTCGCTTTTTCAATTAGTCCTTGCATACGCGCCAGGACCCCAAATCCTTCCACCTTCCCCACAATCCTCGGCTGGTGTCAGAATGCAAATGCAGATTTTTAA
GCCTTGCGGCAGCTTTTGCGCAAACAGATTGTAAACAAATAAATGCCGCCGGCAGACACACAGATACTCGAACATTCGGTGGTCAGTGGCCAGATTGCAGCCTGAAATTCAT
TCAATACTTTATCTCACATGCATTGCACTTCACTTTTGAGCCACGTTAAGTTCAGTCTGCCGAAACCTTTTTCGCACTTAAGTGGTTAAATTGGCTGTGATAAATTGTTGCC
GAGTCAAAGTTGCTGCCGCTGCGGGGCACCTCAAGTCACATGTGGTGATGGTGCGAGTGGGTGGTTGGGTGGCGTGTGCTTTTCTCCTCGGAAAATGCGTGAAAAGTCCACA
AGCACAAAAAGTCTGGGAGGCTGGCTCTCTCTCATCTCATCACGTTCATTGAGTGCCTGTCGCTGACAGTAATCATAACGCAGTCATCAGCTCGTAATATTACGGCCGCCTG
ACATATGACATCTGGGAAAGCGGAAAATGGGGGCTCAGGGACCAAGGAAAAGGTATGCGTACTATTAATATGCCAGTGTGTGAAGGAGCTTTCATTC
(SEQ ID NO: 127)

Exon: 1001..1636

Exon: 1741..2208
Exon: 2283..2681
Start ATG: 1001

Transcript No. : CT37715
ATGCTTCTGAATATTTTAGCTATAATTCTTGTATTCGTAATTTCAAGCCAATCGGAGGCAGTTATTCCTGGATGTGATTACTTCGACACCGTGGATATATCTCATATCCCAA
AGCTAAATGACTCCTATGCGTACGAGGAGCTAATCATTCCTGCTCATCTCACGGGTCTATATACATTTAGGCAACTGGCGGATGGATCACAGGAGCCAGTGAAGAGTCACTT
AAGAGCATGTATCTGCAAACTAAAACCTTGTATCCGATTCTGCTGCCCCCGCAACAAAATGATGCCGAATAGTCGCTGTAGCGATGGACTCACAGAGAATCTGAAGCGGATC
AATCCCTACTTGAAGATCACACTGGAGGATGGAACTATAGGAAAATATTACCTTCTCACCGACATGATCGTCTTAAGATATGAGTTTCGGTATTGCGAGAAGGTTGTTTCGG
TACAGGAGGATCAGTACAAGCTATATGAGGTTAAGATTTCAATCTTTTTGAAGTATCTCAGTATAATTTGATATCTATAGAATGGCAGCTTTATGATTAAACCCGATGTAAA
CTGGACATTGAGCAAGCAGTGGTACTGCCTCCATCCCCGTTTGGAAGATCCGAATTCCATATGGATACTGGAACATTCCATGGTTGGTTGTATTCTCACGATTGCAGTGTAT
CTCTACATAAAGAAACTGCGAAATTTGCTTGGAAAATGTTTCATTTGCTATGTATTTTGTAAGTTTGTGCAGTACCTTATCTGGGCAGGAGGAGATTTGAACTTGTGGAATA
ATATTTGCTCCCTAGCAGGTTAGTTGAATCCATATTAAAGTTTGACTTAAAATATGCGTTTATTCTAAACAGGTTACACTAACTATTTCTTTGCTTTGGCTTCTCACTTTTG
GCTCTCTGTTATGAGTCATCAAATATGGAAAAACTTGAGGTTAATCAATCGGGATGAACGTAGCTATCACTTTCTCATCTACAACATCTACGGTTGGGCACACCTGCTATC
ATGACGGCAATTACATATCTGGTGGATTGGGCCTGGGAGGATAGACCTGACAAACTGAATTGGATTCCCGGTGTTGGCTTATATCGATGCTGGATATGGTCTGCTATGATAT
ACCTATATGGACCGATGCTGATCTTGAGTTTATTCAATGTGGTCACGTTTATCCTGACAGTAAATCACATAATGAAGATTAAGAGTAGCGTTAAGAGTTCTACCCAACAGCA
GCGAAAGTGTATACAGAATAATGAGTAAGCTTTAAGTATTTATAACTTACTATTTCTAATTCCATTCTCTTACTATAGTTTTCTATTATACCTACGACTGTCGGTGATGATG
GGCGTGACTGGGATTTCTGAGGTAATAACTTACTTCGTCAAGCGGCACAAATTTTGGCGACAAGTCCTTCGGGTGCCAAATTTTTTCCATTTGGGTTCCGGTATAGTCGTAT
TTGTGCTATTTATTCTTAAGCGCAGCACATTTCAAATGATAATGGAG
(SEQ ID NO: 128)

Start ATG: 1

MLLNILAIILVFVISSQSEAVIPGCDYFDTVDISHIPKLNDSYAYEELIIPAHLTGLYTFRQLADGSQEPVKSHLRACICKLKPCIRFCCPRNKMMPNSRCSDGLTENLKRI
NPYLKITLEDGTIGKYYLLTDMIVLRYEFRYCEKVVSVQEDQYKLYEVKISIFLKYLSII*
(SEQ ID NO: 129)

Name: mth-like 6
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384207
AGTATATACGAGTATATCCTGTACGACTATATATCAAACCGAAAGTCGAACTGAAAGTGTCCTCTAAGCGGGACTTTCATCGAGGTCGGTGCCCCTCATTATCGTATTCATT
TGTGGGGCTCATTTGATTCATTGTGTCTGTTCGTGTGTGTTGGCGTGTTACTCTGTTAAGTTGGATGCAATTCTGTCGCCGATGGTTGTGGTTTTGTGTCCATGGCCCATCA
ATGGGGATGCTCCATCCATGTGGATGGGCGGTCAGCCAACTAACTTGGAACATTCGTATGTGATGTGTCTGTGGTGGAGGGCGCCGACCTCGAGGCCCCGCTAGCTACTGTAT
CCAATTGAACTGTGATTTAGCTAACTAACTCTAGGTACCGACACCTCGTGTCGGCGTGGAACAATCCGGCCTCTAAGGATATGAATTTCTTTACACTCGCTTACCTAATCGT
ACTGTGGGAATCGGAAGGAAGGTTCTGTGTGCCGTGGGTTTTGGTTTTGGGTGGAGGAAAAGCTTGCTGCTCGTGATATGTGATAGGTTGTGGGAGCAGCCCACTTTGAAG
GACACTACTCTATTCTGGTAAAGCGAATCGTATCCTTGGCGTTTTGCTTACGTCTACTTTGGAATCATAATCGCGTGCTTGCTGTGTGTTGATTATATCGCTACTTGGGGTT
TGCTTAATGTAATTGTTAAAAATTGTACCTAGCCTACAGCTTGCAATCCACCGTTCATGCTCCACATAGCATATATGTATGTATGTAGTTTAGTGGTATGTACAGATCTATT
TACATAAGTAGCTAGTAGAACCTTAAAGCTATTGGGAAACGCAACGAACGACAAACGAAACGAAACTGGAGCCCAATTGAGATCTCTGCTGCAACGATGGCCTCCATTTGGC
ACACTTTGCCATCCTCTCCCGGGCTTTAATCCATTCCACAAGGGACCCATCATCCAATTGCCTATTGCTATAACTATTGCTGGTGACGTGCTCGCCGTTGATAATCACCCCA
TGTGCATGATCTTCTTGAAGGCCTTGCGAAACTCGGGATTGAATATTGTGTAGATAACCGGGTTCACAAAGCTGTTGATGTAGCCCAGCCAGGTGGTCATCATGTAGGCGGT
GAGCCCAGGTCGGCAGTCTTTCTTGAACTTGGCGCACATGGCGTCCATGATGTTGCAGCTGAAGAACGGCAGCCAGCAGAACAGGAAGACACCTGTTTGAGGGTTTGAAGGT
TATATATATGATGGCATATGGATTAGAATATGGCAGAGCGCAAGGACGTGGTCCAAAACGGCTGGGGGAGGTGGCTCAACTTGAACTTACCCAAAACAATGGCCAATGTTTT
GGTGGCCTTGCGCTCCTTTTTGGCCGACGATTTTTCGCGTTTCTTTTTCGAGGCCTTGTGCACCTTGTATATCGTGAATCTGTGCGTCACGCGAAGGATAATAGGCCATGGT
GTGGTATATATTGCATTTTCCACAGGCAAAAAAGGCAAATGATGAAGAGAGGTAGAGAAGGCGAGAAGGAAATGGGGTGGCATTAGCATCGACTGGCCAACAAGAGCCCAAA
TAACTAATAACTGTAATTAGGGGCAATGGACGGCGTTACTCACCTTGACGCCTGCGAGTTCTTCTTATCCTTGCGCGAGGACGTTTTCGATGTAGTCGATAGCGTCGAGTCG
TTGCGTGCCAAAGTCATGGCCTCCTGCACTCCGTACCGCACAAAGGACAATGGCTTCATGGCGACGCTGCGAACGGAAGTCGGTCGGATTCGGAATGAGAACTTAACTGCGG
CATTACGTACTACTAGTCACTATTTTTGGGCTTGCCGGATTGCCTACATTGCCAACATGCGGCTTCTATTTTGGGGGTTTCGAGGATAAAGTGGCAGGTTTCTCGCTAAACA
CTACGTAATAGGTCTGTGATGAAAACGTGAGCTCTTTTTTGCTAAGCAACTTAAATCAAGCGATTGTGCAAGAATGCCACCAAATAATCAATTGATGATTTCATAAGCCTGC
TGATATGTAAACATGCCGACAATAATTATGAAAAGTAGTGGCAAGTGGCAGAACTAATACTGCAAATCACTATTTAAATGTGCAATTGGATTTCAATCTGTTTGAAAAATGT
AATGTGTTAAAAAAAAAAAAAACATGTAATAGTTGACAACAATAACCCGACATATTTAATTAAACTTTGCCATTTGTTTACCAATTAATATTTGGAAGCAAAGATGTTGACA
TTTTTGACGCGGTCAAACTCTTTGTCTGTTTTTATTGAACGAAAAAAGCCATGCGTGTTATACGCATTTATTGTTTTGTTTATAAATAATGACAATTGATGTGTGCAATTTA
AGCAATTTTAATATGCTACGTAATCTAGTGTAAGTAGATAAAAAAAAACATTGCGCAAAAAGAGCACTGCCTATAGCTGAAAAAAGTATTATATTATTCAAAAACTATGTAC
AAAAACTTGAATGGGAATGCTTTATTTTGTCGGGCTGCAAAAACAGTTCATTAAACTTTTAAACTGTTGCGAAAAGTTTCATTTTACCTATGGTACAATCAAAAAGAATGTT
GTTCAATGATTTAGTACCAATGTGGTACAGATATTGGACGGTATTTTGTTTTTTTTTTTTTTCTAAATATTAGCAATTTCTCGAATTTCCGATATAATTGCCGGAACTTTGA
TTTTAAAACAATAACTTGGATGAGGCATTATGTTTTTATCCGTGCACGAACTTCTACTACGCCCCCAACTACTGGGGATGCTATTTCTAAGGTGGCAGGCTTACGGATCTGA
ATGCCAAATACCTGGAGTTATCGACACCGACTGACCTGAGGGCGGGTTCGCCGCGCTTCGGCGAGTCGTCGCCGGTGGGGCGCCGCTGACTGCTGACGCTGGATCGCTGGAG
GGTGGCGCCACTGGGCAGCGGACTGTCCGGCGGCGAGCCGCTGCTCCGCGGAGCGGCTGATGTGGCAGCAGAGGCGGAGGCGGGCGCCACGCCTGCAAGGACATCGTCAACG
TTTGAGGCTGCATAACCAGAATCATGATTACCATTCGGATCGGCGACGACCAGCTGCGGCTGTGTGGTGATTTGGGCCACAACGCTGCTGCGAGTGCGTTATGGCAATGTGG
AGCGATGTGACGGTGGGGTTATCCAATTAATGGTGCCGCGACCATATCCATCAATCAATCAATCAATCGGCGATAGACGACGAGTGGGCGAAGAAGCAGCAGAATCCGGCGT
TGATGAAGGACGAAAAGATGGAAAATCGGGGTGCAACGTCAATTAAAGGCAAGGCAAGATTGAGTTTTAATAAAGTGACGCATTAATTGAATTATTGAGAATTATTTTTCTC
TGCCATTGGCTGTGATTTCTTTTGGCGGGTAAAGTTTCGCGGTCTTGTTTGCCATCCATTTCAGCCCATCGGTTCGGTTCGGTTCGGTTCGGATCGGATCCCATCCCTTTGT
GCATCCAGCCTCCTGTCATTAATGGTTTCATTTCATGTCCTTATTGGATTAAGCAGGCCATCAGGATGCTAATTGGCTGCGCTTAACGCTCGCACTTAATGGCGGTTTCGGC
CAGGAATTGGGAGAAAGATTCGAATGCCGTGGGGCGGGGACATCCGATTGCAAGGATGATGGGCAAGGGATCCTTTCGGCAGGACCTGATTGAAAGTGCAGCACGCTTTTAT
GTCGGCAATTGAGTGCCGCTGTGTCCGGACGAAGCCTGCCATTAGATTGACAGTTTTCCATGGGCATGCACCGTCTGGGGGTCCATTGCTTTGCTGGTCTGCTGGTCTGGAA
GGAAGTGCAGTCGAGACCACTCAGCAGAAACAGGCTTATCTATTACGTGTTTGGGGCACGTGCCAGGGCTGTCAGGGTTGTCAGGTTGCAAGTTGCAAGTTGCAAGCGGACA
AGGATATGGACACGGAAACGAAACTGCTGAGCTCCGAACTTTTGCAGGGCCATCGGAGATAGCCCCCAGCCAAACATTCGTTCTCCAGGCGATGGAGTTTTCTGCGAAATTC
GCCGAAAGTTTCCGCTTTTCATACCCCGTTCACGTAAGATGTAAGTGGGTTTGCAAATGTACATATATGTCCAATCCGACTCCATTGGCAACGTGTTTATATTCGTGAAGTA
ACATTAGCCGCCAAAAGAAGGTCTTACACAGCCAAAAAAAAACTTTGTGAAACGTATATTTCCTTTTTGTATATTCCAAATCTGATTTAAAGCTATCATGTTCTACGCATTT
CTACACAGCACATCAAAGCCCCTAAATGGTATCGTAAAATGAGTACTCACTTGCCCGTCTCCTCGACAACGGTGGCCAGCATGAACTCGGTGGACTTGTCGTTCACGATGAC
GTGGCAGTCGTCGATGTCCGGCGAGATGGCGTTCTCGTCCTCCTCCTCGTTGGAGCCGCTGGCCGTGTTGGTGGCGGCCTCGTCGGGCAGAATCCTACTGGCGTGGCGACTG
CTGTCCAGGGCCGTCTCCGCCAGGCGCCGGGTCTGGGCAATGTTCTCGATGACGCTGCCGCCCGTGAGTTCCGATAGGTGGGGCTTGCGGGCCGCCCCGCTGCTTCCGCGCTC
GACTCCTCAGCGCCTGTGGGTGGGTGGGTACGAGATGTAAAATACAGAAAACAGTATACCAAACATGGGACACAGGATACAGACTACGGACGACGGACCACGGACTACGGGC
TGTGGGACACGGGATGCAGGATGCAGGATGCAGGATGCAGTAGAGAGCGGATTAGGCACTGCCATGCTGCGCTTGCATCCTACGGATATTCGCCCCCAGTCAGCAGCAAATG
TGCACAAACAGAAACAGAAACAGGAAGAGAAACAGAAGCATCCCCAAGCCAGCATTGCAGATTATGGATCTGCGCCAGCAATATTCATTCCACTCGACTCGGAGAAATCCGA

TTTTAATTGCGGTCTTTTAATAATAGTGCTTGAAATTTTTTTGCGCAATGGAAAGAGCGACAGAAAGACCAAAGCGCGGAAAATTCGCTCTAGGAATTAATGTTTGTGGGAA
GCTTGGAGTGTTCAAAGGAATTAACATGTTTGGAATATATCGTATATATACATATATCGCTCGCTACTCACCTTGAAAATGTTCCAGTAGAGAAACACCATGATGATGCAGG
GTATGTAGAAGCTGCTCAGCGAGGAGTAGAGTATAAAGTCGGCGTTGTAGAAGGCGCATACATCGGGTTCGCGGTTGGGCGTGTTGTTGAGGCCCAGAACTATCGGCGAGCC
GATGGCAGCCGATATGGCCCACACCAGCAGTATCGTAAGGCAAACGCGGCGGCTATTTTTGTGCTTGGCGTACTTTATTGGCTGTGTCACAGCGATGTATCTGTGGGATATA
TCCGTGGAATGTGTGCTTGGGTCATATGGGTCTCCGAAATAGGGGCCGCTGTGAACTCACCTGTCTATGGAGATGGCAACTAAGTTGAATATCGATGAAGTAGAGCATATCA
CATCCATGGCTATATAGAAATCACAAACGACGTCAGGCAGGGCCCAGGCTCCATTTACCTGCAAAACAAAAAGAAGGAGGGGGGGGGGGCGTCAGGTAAGAACGAGGATGTG
TATGTCGATGCGGATGAGGAGCACACTTCATTTGTTCGGCACTTAATCGACTCGGGAATGGAGCTCTATTGGAATGCCCATCCTTTGTCGAATATTAATAACAATATTTTCC
CATTTCCCGTGCTGCTGCCACAAATTAGTTTAGACTCTATCTCTTCATCGACGTCTCCGGCAGCATCTATGTTATGGCCCAACCCGAATGCGGACTCCATAAAAACCCCAA
CAAAAACGTCAGGAATTTCCCAAGCCAAACAATTTGTAGCACACAAAGCACATAAAAAAAGGCCGAGTGGGCCCACCAGTTCGGTTGGAACTTAATGTGGATATGTTTGGGTG
AACACCCAACCTGTCGTGGTTGCACTCCCCATAGCCCGTGTACGAGGTATACATAATGTAGCGCCAACCGATTCGGTATTAGAATGGCATCGATAGCTTTTTGGATGTGAAT
ACTATTAGTAAGTAGAGGTAGCTGTCACATAGGTATTTGTTGACTGATCTGCGGGTAAGAGGTATCGCATAGTCGAATACACTCGGCGGAGCCAGTTTCCCAGAATCAGTAG
TTTAACATTTTAAGAAGTCAATATTTGAATAATACTACTACTACATATATTGTACTTATATGCCACCAATGTGTGTGTGCTAACTAAAATGCATGTAGACGGTATCGAATAG
TCGAAGTTAATCGACCCTAACGTACTATCTTGTTTTTCTTCTTCAATCGGTTTTGCGCCTAAATTATACAAACACATTGAAACTTGAAGCGGGGCTCAACGCGTGGGAAGAG
TGTAAAATATTGATACACACGCAACCACACCACACGAAAGTAAAAGTGAAAGTGAAAGTCAAGTGAAAGCAGCATAGAAATGTGCTAAAATAGCTGGGCAGTCGCAGACTTG
GAATTTCTTAGAACTCGGATGAAGACTCCTTAGCCGGAGCTAAGCGAGCATAACAA
(SEQ ID NO: 130)

Exon: 5552..5437
Exon: 5364..5112
Exon: 4605..4307
Exon: 3111..2836
Exon: 1746..1612
Exon: 1423..1323
Exon: 1212..1001
Start ATG: 5495 (Reverse strand: CAT)

Transcript No. : CT37739
TTGCAGGTAAATGGAGCCTGGGCCCTGCCTGACGTCGTTTGTGATTTCTATATAGCCATGGATGTGATATGCTCTACTTCATCGATATTCAACTTAGTTGCCATCTCCATAG
ACAGATACATCGCTGTGACACAGCCAATAAAGTACGCCAAGCACAAAAATAGCCGCCGCGTTTGCCTTACGATACTGCTGGTGTGGGCCATATCGGCTGCCATCGGCTCGCC
GATAGTTCTGGGCCTCAACAACACGCCCAACCGCGAACCCGATGTATGCGCCTTCTACAACGCCGACTTTATACTCTACTCCTCGCTGAGCAGCTTCTACATACCCTGCATC
ATCATGGTGTTTCTCTACTGGAACATTTTCAAGGCGCTGAGGAGTCGAGCGCGGAAGCAGCGGGCGGCCCGCAAGCCCCACCTATCGGAACTCACGGGCGGCAGCGTCATCG
AGAACATTGCCCAGACCCGGCGCCTGGCGGAGACGGCCCTGGACAGCAGTCGCCACGCCAGTAGGATTCTGCCCGACGAGGCCGCCACCAACACGGCCAGCGGCTCCAACGA
GGAGGAGGACGAGAACGCCATCTCGCCGGACATCGACGACTGCCACGTCATCGTGAACGACAAGTCCACCGAGTTCATGCTGGCCACCGTTGTCGAGGAGACGGGCAACAGC
GTTGTGGCCCAAATCACCACACAGCCGCAGCTGGTCGTCGCCGATCCGAATGGTAATCATGATTCTGGTTATGCAGCCTCAAACGTTGACGATGTCCTTGCAGGCGTGGCGC
CCGCCTCCGCCTCTGCTGCCACATCAGCCGCTCCGCGGAGCAGCGGCTCGCCGCCGGACAGTCCGCTGCCCAGTGGCGCCACCCTCCAGCGATCCAGCGTCAGCAGTCAGCG
GCGCCCCACCGGCGACGACTCGCCGAAGCGCGGCGAACCCGCCCTCAGCGTCGCCATGAAGCCATTGTCCTTTGTGCGGTACGGAGTGCAGGAGGCCATGACTTTGGCACGC
AACGACTCGACGCTATCGACTACATCGAAAACGTCCTCGCGCAAGGATAAGAAGAACTCGCAGGCGTCAAGATTCACGATATACAAGGTGCACAAGGCCTCGAAAAAGAAAC
GCGAAAAATCGTCGGCCAAAAAGGAGCGCAAGGCCACCAAAACATTGGCCATTGTTTTGGGTGTCTTCCTGTTCTGCTGGCTGCCGTTCTTCAGCTGCAACATCATGGACGC
CATGTGCGCCAAGTTCAAGAAAGACTGCCGACCTGGGCTCACCGCCTACATGATGACCACCTGGCTGGGCTACATCAACAGCTTTGTGAACCCGGTTATCTACACAATATTC
AATCCCGAGTTTCGCAAGGCCTTCAAGAAGATCATGCACATGGGGTGA
(SEQ ID NO: 131)

Start ATG: 58 (Reverse strand: CAT)

MDVICSTSSIFNLVAISIDRYIAVTQPIKYAKHKNSRRVCLTILLVWAISAAIGSPIVLGLNNTPNREPDVCAFYNADFILYSSLSSFYIPCIIMVFLYWNIFKALRSRARK
QRAARKPHLSELTGGSVIENIAQTRRLAETALDSSRHASRILPDEAATNTASGSNEEEDENAISPDIDDCHVIVNDKSTEFMLATVVEETGNSVVAQITTQPQLVVADPNGN
HDSGYAASNVDDVLAGVAPASASAATSAAPRSSGSPPDSPLPSGATLQRSSVSSQRRPTGDDSPKRGEPALSVAMKPLSFVRYGVQEAMTLARNDSTLSTTSKTSSRKDKKN
SQASRFTIYKVHKASKKKREKSSAKKERKATKTLAIVLGVFLFCWLPFFSCNIMDAMCAKFKKDCRPGLTAYMMTTWLGYINSFVNPVIYTIFNPEFRKAFKKIMHMG*
(SEQ ID NO: 132)

Name: Dopamine Receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384407
TCCATGCCAGGCCTGGATGCACATGGGCCGGATGGGGATCAGGTATATATCGAGAGCGTCGAGAGTTCGGTTTGCGTTTTGCCCACCAAGACGAAACCCAAAAAGGTCGCCT
TTTACGGCAGCAATGGTCAGCGCTACACTTTCCTGTTCAAGGGCATGGAGGATCTCCATCTGGACGAACGCATCATGCAGTTCCTGTCCATCTCGAATGCCATTATGGCCTG
CCGCAGCGACGCACCCGGCAATGGCTGCTATCGTGCCCACCACTATTCGGTTATACCATTGGGACCGCAATCCGGACTGATTAGCTGGGTGGACGGGGTCACACCAGTCTTT
GCGCTCTACAAGAAGTGGCAACAGAGACGGTCCCAAGTTGCTGGGAATGCGGGAGCAGGTGCTGTGGCAAATGTCCCGCGACGTTTCACAGACTTGTTCTACAACAAACTCT
CGCCCCTGTTGGCCAAGCACAATATGCAGGTGAGCGATCCACGACGCCAATGGCCAATATCCGTCCTGCTCCAGGTGCTCGACGAATTGTCACAGGAAACACCCAATGATCT
GCTTGCCCGCGAACTGTGGTGCCAGGCGGGAAATGCCGCCGAGTGGCGGCAATCGGTGCGTCGCTTTGTCCGCTGCATGTCGGTCATGTCGATGATTGGCTATGTCATTGGC
CTGGGCGATCGTCATTTGGACAACGTGCTTATCAACCTGGGAAGCGGCGACATCGTGCATATCGACTACAACGTGTGCTTCGAGAAGGGCCGCACCCTGCGCATCCCCGAGA
AGGTGCCCTTCCGTCTTACCCAGAATCTAGTGCAGGCCATGGGCATCACAGGAATCGAGGTAAGACTATTGCTCATCCTATACCAATCAATTATCTCAACAAATCGCTGCTG
GGGGTGTTTAAGAAACGGTGCTCCTAAAGATAATGTCCACTTTTATTGCTATATCTCCATTTAATTCTACCAAGTTAAAGCGTTGGGTTTCCCAATTCCGTTGATCACTGAG
TGGCCTTCGGTTTGATCTCCTGAGATGAGGATGTGGCCTCCATGCGGTGCTTGCTGTCCGCCGGCAATTGATCCTCGCACTGCCTCGCCGTTCGGCAGCATAATGTTTGGAC
AAAGGCTCGCCGAAAGCCAGAGTTCTTCCACGAGTAGATAATCGGATTGAGAGCAGAGTTGGCAATGGCCAGAGAGAACGTCGTCTTGTAGATCATCGAGCTGCTCTGGCAG
ATGCTAAATAGCTGGGCAATGGCCACACAGAAGTACGGTAGCCAACAGAGCGTGAAGCAGCCCATGATGAAGACCACAATCTGGACGCTCTTCCAGTCTGGATGGAGCAGCA
GATTTCGCATGGTGGTGGCCTCGTCCGTGTTGTAGACCACCGACTGTCGCAATCGCAGCGCCTGCTTGGATGCCTCGCGCATGATGCGCCAGTAGACGAGGAACATACAGAT
CCAGATGATCACAAAACCCGGCGTTATCACTCCGGCAATGTAACCAGGCGCGAGAACTTCGTCGAATTCGCACGCCTGTGCATCCGGCCATCGATTCCAGAACACTGGCAGC
AGAGCCACCAATGCACCCAAGCACCAATTGAATATGATGATGCTATATGCCACTCGACGGGTCATGTATCTACGGTTTCAAAGGGTTTCACATATGTAGCCATTGAATTAGG
ATTTATCGCGTTCGCTTAATACCTTACCTTCTATAGTGCAGAGCATAGACGACTGCTATATACCGATCCACCGCAATCGAGATCAGAGTCAACATGGACACGCAGCAGGCGC
AGATGAGCAGGAAGAAGCGCAGCAGGCAGGGACCTCTCATGGCGCCAATGTCCGAGCCCATGTAAAATACCAGATGATAGGGCAGTGCCAGTCCCACACAGAAGTCGGAGAC
GGCCAGCGACAGGATGAACAGATTGGAGATGACCGATCGCAGGTGGCGGCAGGTGCGCACCGCCACAATGGTCAGAATATTGCGCCCCAGGATAAGGACAAACAAGAAGGCA
TTGATGGCTAGCCACAGCAGATGGCCGGCACCGAAACTGGTCGCCTGGAGGGTTATTTCCGTTGGCGTTTCCGTGGCGTTCTGGTGCTGCCACAGGAACGGACGACACCCA
ACGCTTCCGGGGAGCTTCGATAATCCCAAGCGGGCGATGAGCTGCTGGTAAATGCCATGAGCCACGATGGTAAACGAAGCGCACGCGACGCCAGTGGTTGGCTATCTGGCAA
CCGTGGGAGCTGGAGTTGTGCTGACTTTTGGCCAGGACTTCGGCTCGGCTCGGATCGGATTCAGATTCAATTGGGTTTCGGCTCGGACCGTGGGCGTCGTTTGTTTACGTTT

CGCTGGTGGCTTTACGACATCGTCATCGCGCCAGTTAACTGATAGCGCGCCCAGATCCACCGTGAGAACTGCAATCCGTTGATAAGCCCTCCACTTTGGCCCAACCACCATT
CACCACCCTCATACCTAACCACTCACCCACCCATCGAGTCAATCCAACCACTCCAACCAGCCAAACTATCCATTGCCCCCGAGTATTTCGAGTGTCCGGACTCCCGGGCCCA
ATGCATGGCCAAACAGATCGCTGGGGGTTTTAGCAGATTAGCCCAGCTCCGCGGACCAAAATAGATTGATAAAAGGCCAGGGCCAGCAACTGGGATACGTGGTAAACAGATA
CATCATTGATATTCCCAAATTTCCTCGATACATTGCAAGCGAAGAAACTATATGTAAATAGTGATATCATCATAATTGTTATAAATCCGAGACCAACAGTACAACCAGTACT
ATTAGGATTGACATGGTTATCCACTCTTGGTTCTTTTAGAATTAGAAATCAGGAGAATATATAGTATATATAGGTACAACTCTCTGATCTAATCCCATATTTGAATTATCTT
TAAGATGAGATATTAAAGATTTCATGGATCAGAACAGACCAGGAGTTTGCTAAGGATG
(SEQ ID NO: 133)

Exon: 2186..1708
Exon: 1637..1001
Start ATG: 2186 (Reverse strand: CAT)

Transcript No. : CT38338
ATGGCATTTACCAGCAGCTCATCGCCCGCTTGGGATTATCGAAGCTCCCCGGAAGCGTTGGGTGTCGTCCCGTTCCTGTGGCAGCACCAGAACGCCACGGAAACGCCAACGG
AAATAACCCTCCAGGCGACCAGTTTCGGTGCCGGCCATCTGCTGTGGCTAGCCATCAATGCCTTCTTGTTTGTCCTTATCCTGGGCGGCAATATTCTGACCATTGTGGCGGT
GCGCACCTGCCGCCACCTGCGATCGGTCATCTCCAATCTGTTCATCCTGTCGCTGGCCGTCTCCGACTTCTGTGTGGGACTGGCACTGCCCTATCATCTGGTATTTTACATG
GGCTCGGACATTGGCGCCATGAGAGGTCCCTGCCTGCTGCGCTTCTTCCTGCTCATCTGCGCCTGCTGCGTGTCCATGTTGACTCTGATCTCGATTGCGGTGGATCGGTATA
TAGCAGTCGTCTATGCTCTGCACTATAGAAGATACATGACCCGTCGAGTGGCATATAGCATCATCATATTCAATTGGTGCTTGGGTGCATTGGTGGCTCTGCTGCCAGTGTT
CTGGAATCGATGGCCGGATGCACAGGCGTGCGAATTCGACGAAGTTCTCGCGCCTGGTTACATTGCCGGAGTGATAACGCCGGGTTTTGTGATCATCTGGATCTGTATGTTC
CTCGTCTACTGGCGCATCATGCGCGAGGCATCCAAGCAGGCGCTGCGATTGCGACAGTCGGTGGTCTACAACACGGACGAGGCCACCACCATGCGAAATCTGCTGCTCCATC
CAGACTGGAAGAGCGTCCAGATTGTGGTCTTCATCATGGGCTGCTTCACGCTCTGTTGGCTACCGTACTTCTGTGTGGCCATTGCCCAGCTATTTAGCATCTGCCAGAGCAG
CTCGATGATCTACAAGACGACGTTCTCTCTGGCCATTGCCAACTCTGCTCTCAATCCGATTATCTACTCGTGGAAGAACTCTGGCTTTCGGCGAGCCTTTGTCCAAACATTA
TGCTGCCGAACGGCGAGGCAGTGCGAGGATCAATTGCCGGCGGACAGCAAGCACCGCATGGAGGCCACATCCTCATCTCAGGAGATCAAACCGAAGGCCACTCAGTGA
(SEQ ID NO: 134)

Start ATG: 1 (Reverse strand: CAT)

MAFTSSSSPAWDYRSSPEALGVVPFLWQHQNATETPTEITLQATSFGAGHLLWLAINAFLFVLILGGNILTIVAVRTCRHLRSVISNLFILSLAVSDFCVGLALPYHLVFYM
GSDIGAMRGPCLLRFFLLICACCVSMLTLISIAVDRYIAVVYALHYRRYMTRRVAYSIIIFNWCLGALVALLPVFWNRWPDAQACEFDEVLAPGYIAGVITPGFVIIWICMF
LVYWRIMREASKQALRLRQSVVYNTDEATTMRNLLLHPDWKSVQIVVFIMGCFTLCWLPYFCVAIAQLFSICQSSSMIYKTTFSLAIANSALNPIIYSWKNSGFRRAFVQTL
CCRTARQCEDQLPADSKHRMEATSSSQEIKPKATQ*
(SEQ ID NO: 135)

Name: Adrenergic receptor- like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384675
GCCTAAACGATAAGATTAGAAAATGGAAATGGAAATGAAGACAGCCACCTGACGCAATCTCACATGCCAGAAATTTGTCGTTTGCCATCCAAGGCCCCGATTTCCGAAATCG
AAAATCCTTGATATCCAACTGGACATGACCAGTTTCCGGAGGAATGCACTCCAGTTACACAACTGGCACAAACATTTGTTTGCTGTTTTCTGACCAAAAACAAGCCACACTT
ATGCTGGCACTTTTTTTGTTGACAGAAATATGCAAGCAGACACAAGTAAATATTTATTTGATTATTTATTTTAAAGAAATGTGCAAGCTCACCGAATAAATATTATATTTTAA
AATCAATAAGTAGAAACCTAAACAAATACAAGCTGGTGAAGCGATTGCTCTGCAAATTCCACATATTTAAACTATTAATTTACAAAAATAAATGTGGCGGTCAATTTCCAGC
TGAACAGTTCTCAACAATTAATGAAAATAATTGCCACTGATAATAATAGTTTTTTTTATTATTCGATTACATACTTTAAATATTTTTTTAGTTTCCCAATTAGAAAACTTTT
TTTTAACAAAATATTTAATATTCCGTAACAGTATTTATGCTTAATTACCAGATGACAGTTTTTATTTCTTTTGACTATACTTACATTATTTTGAAAAAATAGCAATTATTTT
TAGACTTCTGCCACGAATGCCTTTAATTAATAGCTGATGCAATCGCTTAAACTCAATTAAGTTACACACGCAACCCACATCTTCTGTGTAAAAAAAAATCAATGTATTCGAC
CATAAATAATTCCGCTGCAATCAGTGCAACAGTATTCAGTCCTGGAACTTTAATAGCGCCAAGCTGTTGGAAAACCAATTAAATTCCCTTTCAAATGGAGCGTAAAACGCAA
AATACTTACGTATGATGGGGGCGCCACACACTCGCACACGCATTTATGAGCTCGGCAAAACCCATTTGATTGACTAACGAAACTTATTTGCCACTTTATTTCAGCGGACGCA
GTGCAGCAGTTCCACGACGACCATGAGCGACCAGATTGGCAACCCCAATGCAACATTCAGTGGCAGTGGCAGTGGCAGTGGCACCAACGTCGCCTCCATCGCCGAAAGTGTG
GCGGAGAGCGGACCGGACTTCGATGCACTTCGGGCCGCCTGCGAAACGCGGCTAAATGCCAGCGGTCAACTGGCGGGTGAGTCGTGCACTGCGAGAAAATATATCACACAGC
ACATAACTATTAAAGTTTATGCAGTTATTTTATTTGATGCAGTTGTAAAACGCTCACCAGTAAATTTCATTTAAATTTTAGATGTTTTTATAAACATAAATCAAAATAAATT
TTTTTGCACTGTTTACTTACATTCCCTTTCTTTATGCCAGTTATTTGTCTTTACTTAAGCGAAAGAAAAAATTCACTCTGAAATTGTAGCACTAGGGAAATATAACAGTTT
TTGGTATTTCACTACATATTTTTGCCCGTGCTTCTGGCTCTGGCCAACTGCAATTTCGGCGTCCATTATGCACGAGTAACTTTAAATTGCTGACCTAACCACTCCCCTTTCC
TTTTGGCCCCTTGGCAAACAGGCTCCGGAGGACCAGGCGCAGAAGCAGGAACCCACTGTGCCGGCACCTTTGATGGCTGGCTTGTTGGCCGGATACGGCTGTCGGCACTTC
CGCCTACGAACTCTGCCCGGACTTCATCACGGGATTCGATCCAGCAAGTAAGTGCGCCGGGGCAGGAGCTGAATTGTTATTAAATTCCACAAAGCTACATCCGTTCGGGG
TCAGCTGGGGATCTGCATATGACACGAAAAAATATATATATATACATACATATATCACATGTATGTATGTTACATATAGGTTAAATAGCATGCCAAAGGGAAATGTGTTTGC
ACGAGGCAATGCGGTTGGATAAGTTGGTTGGGTTGTTTGGTAGGTGGAGTGTGCCAGTCTACCCTCTAATTGGGCTCATCCCCGAGCTCATCGTTATATACACATAAATGGA
ACAGAAAATACGCTCCAAGCCAGGCCAAATCAATTGGAAACTTGTTATGTTTATGCTTTGATTTCGTGTCGTATGGAAACATCAAGTATTCACAAAACCACACGGGCTAAAG
GCAAAGCTCCTTGGAGCTCCCTGGGCCACTAAATTATCCAAAAGAGCTCGACAAAATGCTAAAATGTTTCGCCTGATTTGGCACAATTTGCCAGAGTCTAACCGAAATCAAA
TGTAAGGCAGAAAACACGGCCAAATTGATTGAGGCTGCTGGGCAGAGTAGAAGTCAAATATGTGTGTGGGTATTTAAAGATGCTGGAGCACGAAGAAAGGGGATTTAAATTT
TATCATGTTTGTGAATTTCATATTTTTTTAATCATTAATCAGTGTTCTTGGCAAGGTGTTGATGACGGTCATAATAACATAACTGAAGAAGAAAACCCTTTGAATAGTTTCA
CCAACTCATTGTCGTATCATTAACATATCATAAATAAAAATATCAAAGGTGACTTGAATTATTTCTCAACTTTCTGCCACACTGCGTGTATTTCATTCAATTATAAATATAT
ATATATATAAATATAAATATAAATTAATCAACTGATTGGCCAAACGCACAAAGAAAACTAATTAATTATGCTTGGCATTAAAGCAAACGTTTAAAAACAAAGCGCACGCTCT
AGAAAATCGTCATAAATGTGGATAAATATTTGAACCAATTTATGGCATAAATCATGCGCTGCATGATAATTTGCATTAATGGTATACAATATGTAGCCAGCCAACTTAAACC
TCAACCTCGTTTTGGCCATGACCAGGGCAACGAAAATAGCGAGAGTTGAATGACTTTTTTTGTTATGACAGAGCGGAGAAAAACAAGGCGAGCGTCAAGAAAGTGCTACTTT
TGCTACATTTAGCTTGTGGTTAGCATGGCGCCACCCCCTTTTTCATCCTTTTTGCATTTTTTTTTTTGCCTTACGGGGTCGCCGCCAGCTTGTACAAGATTATGTGCTTTGG
CACTGCGCCGCCATTAACTTTTATGGCACTTCCGGTCGCCGGATGTCGATGTGGTGAGCTGTGGTCGAGAACTCCCTCACCCCCCGGCTACTTATCTGCCACACAGCGCGTA
TACGTGATTTTCTCGGTCTGCGTGCCAAACGCCGATAAATCGGATTCTTGGCCAAAGATCATCGGTCGAAAAAAAAAAGAAAACGGAAGAGCCGGAAAAACGAACGACCTGAT
AAAACTCACCTTCGAAGTTCCTAACAAAAATTCTCCAGTAATTTTTCTACAAGTGCCACTCGGTAAATTCATTGCAAACTTGGTAATCTACAATCGAATCGAATGGAAAGT
TTTTGTCACGAACTTCGAATTGATTCGACTTAAACACACACATGACCCAAACAACAGAATGATTATGCAATTTTCAAATTGATATGCTTGCAATTCGAATGCGAACCGCAAT
TCTTTTGCGGTTAACTTGGTGGCAGATGCAATAAATAATGGTTAATTCAATTGAATGCTCGGGCCGACCGCAAAATGTAAATATGAGCCGGAATATTTATGCAGTAGTCAAT
CATGAAATCACACATATTTAGCAAGTGTAAATATTTGTGTGGTGCACAGGATGTGAAAATGTTTGCTAGAAAAAACGGTAAAGAATTAATCGTATATACAATATTTATTGAG
GAGCTGTAATAAGCAGAGTAAACAGCATGGTTAATTATTTTTTATTTATCGCAAACTGACATTGCTTTAATTATGCAGGATATGCCCACAAGGAATGCGGTCTCGATGGCGA
GTGGTTCAAGCATCCGCTGACCAATAAAACATGGTCCAACTACACAACCTGCGTAAATCTCGAAGACCTCAACGTAAGCCGATCCCATACATAAATATCCTATATATCCTTT
ATGTAATTAATTAATCGCAACACTATATATCCAGTGGAGGCACACAGTGAATCTGATCTCCGAGGTTGGCTACGGCACCTCACTGCTGGCCATTCTGCTGTCGTTGGCCATA
TTGGGTTATTTCAAGTGAGTATGGCCAAGGGCTTTGCACTCCTTCATCTTGGCCACAAGTTGACAGCTTCGTTGGGAGATGGCAAAAACACTTAGGGGTAATGAGGCACGCT
GGTGACTTATATTGCGACTTATATTATGACGACTGGCGCCATAAGTGAAATTTACTGATGGCAAACAAATGCGACAGTGGCTTAATCAACTGCCAACCTGGATGCTGGATGC

CCAAATGCTCATTGCACTCATTGCTCATTTAATGCGGCACAATTGACGCCGGAAATGGGAAAATACATTGTGTCCCCATCTCCTGGTGACTCGAAGCAGGATATCGATTAAT
TTCCCAGTGCCTGCAGACCAATTAGTGCTCTGGGTTTGTTTATCGCACATTCGAATGGAACTGGGGCTCAAGGATAAAGGATAATGGATGTGCGGACTGAGGGGCGAGATCC
CATAGACGGTTCGGCCAACTGACAGAACTGTTTTAAATTGTGTGCCAAAGGCAAATAAAATGGGTCCAGATACGTGTGGGAATTTGGTCTGCCCTAAAAAAGTGAATAATTA
AAAGTTTAATGACAGAGCAATGGGCAACACCAGCCCCGGAAAAACGCATTCCTCAAACAGCTGTAATGAAAATACCAGAGTTTTCATTCAAATAAAATATAAGAGTATAAAG
GAAATTGAGTCGTTGTGAGCGGCCAAAAATAAATAAAGAAGTATTATTATAAAATATACCATCAGTATATTCTCAGTCAATAATTGTATTGCTTAAAGAAAGTTTAGAAGTG
AAAAGTAGTATATAAAATAATTTATTTAGACAGAAGTAAATTAAGTAAAGTAGAGAAGGGATTAACGAAAACAATCTAACCTATTGAATTAAATTGCTGAAAAGAAATCAAA
CTAGATTCAGTTTGGCCAGATGTCAGCAATCAATGCAATTGAGCAAAGGAGAAACGGAAAATGGGAAAATGGACAAATGAACTTTGTCCTTTTGGTCGCAGGACTTTTTCTG
CCTCTTTCGGCCCTATCCAAGCTGCCTCACATGCCCCTGGCCCCGAATGCAAATGAAGCCATAGACAAACTTGCAAATTTGGCCATAAAAACTTGCCACAAATTGCCGGCCA
AAGTAATTCGATTCGCATATTTGTCTTTCGATTGAACAGATCCCTGAAGTGCGCCCGCATCACGCTGCACATGAATCTGTTCGCCTCGTTCGCTGCCAATAACTCGTTGTGG
CTGGTCTGGTACCTGCTGGTCATGCCGAATTCGGAGCTACTGCATCAGAGTCCGGTAAGTGTGCACACCTGGCGAATGCGTAATGGTGAATCAACCTCGGCTTCCGATTTCC
ATATTCCATACTCCGTGCGAACACCCAGATGCGCTGCGTTGCCCTGCACATAACGCTACACTACTTCCTCCTGTCCAATTACTCCTGGATGCTCTGCGAGGGATTCTATCTG
CACACCGTCCTCGTCGCCGCATTCATTTCCGAGAAGAGGCTGGTCAAATGGCTCATCGCATTCGGCTGGGGCTCCCCGGCCATCGTCATATTCGTCTATAGCATGGCTCGCG
GTCTGGGCGGCACGCCCGAGGACAATCGTCAGTAAGTGAATAATGAATGGGTACAGCTTACACCGAAAGAAAATGAGGAATGCCAATAGTTATCAAAAGTGACCGACAACAG
ACAAGGTGTAACCATATTGTGTTGTCTAGAAGTATCCTAGTTTAATGAAATCCCACAGACTTGAGTATCACAGATTGATATCAAGATATTTCTGTGTAGCTAAGATATTCAG
TCTGTCGAGCTCTTGACCAAATGCTAATCAAATCAATCAGCAATTCACCCACTGACTCATCTCGCATTCCTTCGCCAACTTGCAGCTGCTGGATGAACCAAACCAACTACCA
AAACATTCTTATGGTGCCTGTGTGTATCTCCATGTTCCTGAACCTCCTGTTCCTGTGCAACATTGTGCGAGTGGTCCTTTTGAAACTGAATGCCCCGGCCAGTATTCAGGGC
AGCTGCGGTCCATCGCGAACGGTTTTGCAAGCATTTCGGTAAGTTTTCCCGTCCAAAACAAAAGAAAAAAAAAAAGAACTGTTGGGGGGGAGATTGAAACACTTATAAGCGAA
GCAGAGAACTCTCAACTGGCCCACGGGGCGAATGCGTAATTATATTTTGTTTCCCTTGTATTTAGACTCTTTCCCATCGGCCATCAGGTCTTTTATTGTTTATTCGCTCTGC
ACTCTGGAAGTTGGGCTCATTTTACATTTTGAATTACGCATACGCCGCGTTGACTTGCATGTTAATACACAGCTGAAAAACGAAAACAAGGCATTATAATTCACCCTCTCAT
CCTGTTTTAGGGCAACGCTGCTTTTGGTTCCTCTGCTCGGCCTCCAGTACATCCTTACGCCCTTTCGTCCTGCACCCAAACATCCCTGGGAGAATACATACGAAATCATCTC
GGCGTTTACGGCCTCATTTCAAGTAAGTAAAATGTTTTCGTTTCGTTTCTGTTCTCGTTTCGTTTTTTTTTCGTTTTTTTTTTTTGCTTTTTTCTTTCCGGCTTTTGTTTTG
CACGCTCTCATAAAATCACTTTACGACTCCGAGCTGGCTAAAAAGTTATGATGTTAGTTCGCCTCATGGTGCCATCCACCACTGGGAATGCAAATAAGCCAGCATTACTGGA
ATACCGAATACCAAATACCAGGACTAGCAGCTCATTAGATCGGGGCTCTTTTCAAATCAATAATGAAAATACACGCCATGTGAAATACCCACCATTGGCATTTCACTTGATC
TCAGTTTATTAAATTGCTGGCCGAGATTGGAGCGAAGTTCCCCCGTTGGAAAAATGTAGCAAGTTGAGCATGAAATGCACTCTCGCACACAGGAACTCGTAAAAGCAACTCG
AGCTTGTCCCTCGGCCAACTCAGTCTCGATAGACCAGCATTTACACATGGTTAACACACATACATATATTTCTTGGCCGCAACTTAAAGACAACGACGGCAAACTTCATCCA
GCTACGTTGCCTTCTTGTTTGTTGAATACCCAAAGAATCTTTCAGTTCATGTTTGCAGTACACAGATAAAATAATTTGCATATATTTCGTGTTTTTGTGCACGAAATGGTT
TTTCATTACAGACTGTTGATATAACAGTTGCCATGGAGGTAGAGGTGCCTAAAAGTATGTTGCAGCTTAGATGGCTTTCCTTTACAAAATTTCAAAAGTAATTTCTAAGGCA
CCCCCTAGCTGTTATTCGAAAATCTTTTAGTTCGATCTAAAACCCAACTCAATCGGTTCAAAAAGGTAGCCCATTGAATTCATAAACAGCCCTAAAACAACAGGAAAATCCA
TTATCTTGCGAGTAGACATAAAATCATCTTGGGTGCCACTAAGCTGCAGTTGAGCTCTTGACTATTTTCTATAACTTGGCAAACTGAACATTGTCACCTCGGTCTGAGCCTT
TTTTCATCATTGCCAATCGATAGGTGAAATCTTTTATTCGACTGCTAAAAGTTTAAAGCCAGTCGACGGGTACTTTTCCTTTGCATTTTCCTCTTGTCGTATTTATTCACTT
TTTTTTATGTCGCTGTTTTCCGTTGGCTTATTTGCTGCAAAGA
(SEQ ID NO: 136)

Exon: 1001..1196
Exon: 1590..1727
Exon: 3775..3881
Exon: 3955..4046
Exon: 5156..5318
Exon: 5405..5631
Exon: 5910..6086
Exon: 6395..6659
Start ATG: 1031

Transcript No. : CT38445
CGGACGCAGTGCAGCAGTTCCACGACGACCATGAGCGACCAGATTGGCAACCCCAATGCAACATTCAGTGGCAGTGGCAGTGGCAGTGGCACCAACGTCGCCTCCATCGCCG
AAAGTGTGGCGGAGAGCGGACCGGACTTCGATGCACTTCGGGCCGCCTGCGAAACGCGGCTAAATGCCAGCGGTCAACTGGCGGGCTCCGGAGGACCAGGCGCAGAAGCAGG
AACCCACTGTGCCGGCACCTTTGATGGCTGGCTTTGTTGGCCGGATACGGCTGTCGGCACTTCCGCCTACGAACTCTGCCCGGACTTCATCACGGGATTCGATCCAGCAAGA
TATGCCCACAAGGAATGCGGTCTCGATGGCGAGTGGTTCAAGCATCCGCTGACCAATAAAACATGGTCCAACTACACAACCTGCGTAAATCTCGAAGACCTCAACTGGAGGC
ACACAGTGAATCTGATCTCCGAGGTTGGCTACGGCACCTCACTGCTGGCCATTCTGCTGTCGTTGGCCATATTGGGTTATTTCAATAATTCGATTCGCATATTTGTCTTTCG
ATTGAACAGATCCCTGAAGTGCGCCCGCATCACGCTGCACATGAATCTGTTCGCCTCGTTCGCTGCCAATAACTCGTTGTGGCTGGTCTGGTACCTGCTGGTCATGCCGAAT
TCGGAGCTACTGCATCAGAGTCCGATGCGCTGCGTTGCCCTGCACATAACGCTACACTACTTCCTCCTGTCCAATTACTCCTGGATGCTCTGCGAGGGATTCTATCTGCACA
CCGTCCTCGTCGCCGCATTCATTTCCGAGAAGAGGCTGGTCAAATGGCTCATCGCATTCGGCTGGGGCTCCCCGGCCATCGTCATATTCGTCTATAGCATGGCTCGCGGTCT
GGGCGGCACGCCCGAGGACAATCGTCACTGCTGGATGAACCAAACCAACTACCAAAACATTCTTATGGTGCCTGTGTGTATCTCCATGTTCCTGAACCTCCTGTTCCTGTGC
AACATTGTGCGAGTGGTCCTTTTGAAACTGAATGCCCCGGCCAGTATTCAGGGCAGCTGCGGTCCATCGCGAACGGTTTTGCAAGCATTTCGGGCAACGCTGCTTTTGGTTC
CTCTGCTCGGCCTCCAGTACATCCTTACGCCCTTTCGTCCTGCACCCAAACATCCCTGGGAGAATACATACGAAATCATCTCGGCGTTTACGGCCTCATTTCAAGTAAGTAA
AATGTTTTCGTTTCGTTTCTGTTCTCGTTTCGTTTTTTTTTCGTTTTTTTTTTTTGCTTTTTTCTTTCCGGCTTTTGTTTTGCACGCTCTCATAAAATCACTTTACGACTCC
GAGCTGGCTAAAAAGTTATGA
(SEQ ID NO: 137)

Start ATG: 31

MSDQIGNPNATFSGSGSGSGTNVASIAESVAESGPDFDALRAACETRLNASGQLAGSGGPGAEAGTHCAGTFDGWLCWPDTAVGTSAYELCPDFITGFDPARYAHKECGLDG
EWFKHPLTNKTWSNYTTCVNLEDLNWRHTVNLISEVGYGTSLLAILLSLAILGYFNNSIRIFVFRLNRSLKCARITLHMNLFASFAANNSLWLVWYLLVMPNSELLHQSPMR
CVALHITLHYFLLSNYSWMLCEGFYLHTVLVAAFISEKRLVKWLIAFGWGSPAIVIFVYSMARGLGGTPEDNRHCWMNQTNYQNILMVPVCISMFLNLLFLCNIVRVVLLKL
NAPASIQGSCGPSRTVLQAFRATLLLVPLLGLQYILTPFRPAPKHPWENTYEIISAFTASFQVSKMFSFRFCSRFVFFSFFFFAFFFPAFVLHALIKSLYDSELAKKL*
(SEQ ID NO: 138)

Name: Calcitonin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384816
ATTCGTTCGGTATTTTCGCATTTGGCATGTTTGTGGCTAAGATATCATCAATTCAACACAATGATCAACAATATTCAACGACAAACACTCGACTTGGGACCCTGGATGCTTC
ATCGGCGGCGCTCGGGAAACTGAATTTGACCCACTAAGGGAACTAGGCAAAATTGCAACGCTCTGTTGGTAAGCCCACCAAATCGAAGCGTTCAGGGGCGGCCAAAAGGAAA
ATGAATCATTCTACCGATTTGAACAACGATCGAAACTTCACACTAGCCTATCATTTCAGTATCATAGTTGTATGTTGTATGGCTAAGAACTACTGGAAATGTCTTTTTGGGC
TTATTTGCTTTTTGGATTTTAACATATTTGCTGTGTTTTATATTCGTTTAGTGGGCCTTGGTTTTTCGTTTCTTAAGCGCAATTCGCAAAGCGTGCGAAGATTGAAGATTTT

CCGAATTTTGAATTCTCATATTTCGGCCCGTCGTCAAGTTTGGCAAATTGTTTAATAAACGCTACACGATTTCATTTTGTAAAATTCTAAATGCTACTTAATATGGACTTTG
AAAAAGGTTTATACTTATGCAATCTTTAGATAATAATCCTGCTGCTCGAGTAGCATATGATTTCGTATATTGTTTTTGTTGTAGCAAGTTCGAATTGTAAACCGGAAACGG
AAATGTTGCTCTTGCCTAGTGAATTGTCTCTTGTCTACGAATTAGAATTAGAATTAGATCGCTGCTACATATGTATATGTATATGGCTAGAGTTAAGTACAAGTATATATAT
ACGTGGTAAGCTCATAAATACACAATAATAATCGTCAATCAATTCAATCAATCAATCGGTGAATCAATCATAATTGGATACGCTAAAAACTCTTAGCTTAGCTTAGGTCTAG
AAATATGAATATGTCTTTCAACGTGTCTGTGGTTTTGGTGTTTTTGGTGTTTCTGTGTGGTTTGCTGTATGGTGGCTGAGTGGTTGAGTGGTTGCGTTTTTAGTTTTCAGTT
TTTCTGTGGCTGTGACTGTGGGGTGGTGCTTTCATTGGCATGTCGCGTAGCGCCAAACGGCCCGTAAATATGCAACTATATGGCAAACAGGCAAGTAAGTGTTTATATATAT
TGAGAGAGAGAGATATATAGAGAGTTAAGGAAATGGCTCAATGCAATTACAACCAATTATTTACTTGCAGGACTTTGTAGTTCTTCTCAAAGGGATTACGGAGGATGGGTTT
GGGAGGGGGATTTTTACGGATAATTATTGTATTTTTTTTTGAGGAGAATGAAAATCTAAATCTAACTCTAGTCATCTGGGTCCAACCCGATTAGCTCTGCTTGCAAACGGCC
CTTAGTCTTACAACAAATGGTTAAGCAAAATACTCTCAAAGCGATACGGAATTGCGGCGACATACTGGTCATGATCAGAATTTTCCAAAACGAATTCAATATGCCGATCCAG
ACGACGGCGAAGTTGATAAGAGTACTTTGGATAACATCGCCCGTGACCACCTCGTACAGACGCAACAGAATCCATGGCAGCCAGGACACCCAGAATGCAAAGACTAATGCGA
ATGACATCATATGACTAGGGTTCGATAAATTTTCAGCCAGAGCAGTTGCATATTCTTTATCGTGTATCGGCTCGCCGGACCTAATCTTACGCATCATTACAAAGATATAGCC
GTAGTTGTGTACGATTGAAATGAGGCTGGGACCTAAGACTAATATCGCCAAGGTGGCGCTATATGCTGCCATATTCCCCCAGTCCAGCATGCAAATGTGCGTCATGTTGTTC
TCGATGGGCATCGAGTAGCCCAGGATTGGGGGACAGCACAGCAGGGCCGCCGATATCCAAGTGAACACCATCCAGCATTGGCATCTGCACGATAAAGGTCACAAGAGATCAA
AGATTAATCCAGTTTTAATCCAGATGAGGAGGAATTCGGGGTTCCATTGGCGCTACTCACCTCGTTTTCGTCTGCACCGTTTCATATCTGAGTGGCTTCCGCACCGCAAGAT
AGCGATCCACTGATATCCACATGAAGGTGTATACCGATACCGCCCACAGAGTGACCTCCAGATAGCCGGTGAAGCGACACACGATGTCGCCGTACATCCATTCTCCGGTCAA
AGCGGGATACACGGAGAAGGGCACCACCAGGAGTCCGCAGAGCAGATCGGCGATGGCCAGGGACAAGAGGTAGTAGTTGATCACCTCTGTGGGTCCTGCGATGACAGCAATA
TCGTTAGAAAACGAAGTAGAGGTGTATTCAAAAATAATAGGCAAATGCAAACAATAATATTTAAATGGATCATTCAATAATTTGGGCGAGTCTATTTGTACGACATTCACTA
GAACTATTTAGAAACCTAAGGTTCTTAAATTTATCCAATAACAATCAATTTATAGGATAAAAGTTTAGGTACAAAAAGGGGAAAGAAATAATATGCTTAAGCACAATTGGTT
TTTGAATTTTTTCAAAGCAATTAAGACAACAAGAAAACAACACATTTGAATTTTTGGGGAGGATGCCGTGTTACTTCATTCCTTTCACTTTCACATCAAAATCATGGATGCA
TTTTAAGTATAAACATTGGACTTAAGTGGTCAAAATAAATTATATTCAACACATATGTAGGTTACTTATGAAAATATTATGCTTGAGGCCAAACCCAACCGCGAGATTCGAG
ATGTTCCTGTGGAAAGTGTTATGTGAGTGTGATATTCGTGCAACGTACCCTTGAAGTTGGCATAGGTGGCGATAATGAGGAGGTTGGACAGGACGATGGCCACGCCCAGTAT
CGATATGAGCACCGCCTTGGTGAGGGCCTCCACCTTGGAATTATCCTGTAGGTAGCTCATTTCCTGCATTTTTACGGGCTGGCACTGTCGGTTGAAATGGGCGCTCCTGTGC
TCGTTGTATACGCTCCCTTAGCTCCTTGTCCGTTTCCCTCCTCCTTGGCTCCTCCTGCTTTCACCTGCTTTCGCTGCTGCTTTCACCCGCTTTCTCCGCACCTTCTTGGCGG
GCTGATTGTTATGCAAAGGCCAGTGGACGTGGACGTGACCGTACTTTACTTCCGGTCATGGCGGACGCCCATGCAATTGTCCTGTGCCTTCGGTTGTGTGTGTCGGTGTGTC
TGTGAGTGTCTATGTGTGCTTAGGTGAGTGTGTCGATGTGTGAGTGTGTGTGTGTGTCTGGGTGCGAGTGTTGGTGGCGTCCTTTATTGTGCAACGAGGCTTGCACACGCGA
AGGTCGTCGGAGTCCTTGGCGCTGTTTTTCGTCCTTTCTCCAACGTTGGCGCCAAATTTGTATAAATTTTCTACTGCCTACTTTTATTTGGGCTGCAACTTCTACTACTGC
CCACCTTTTGGCGCCGAGCCAAGCATTTGATTGTTACCCTCACAATCCTGGCGTCGTGGTTAAGGTTAAGGATAAGGTTGTGCTGGTGTGTGTGTGTCTATGTGTGTGTCTG
AAGCTCGTTTAAAATGCTGCTTGGGCTAAAATCTTACACGTAAAATCTAGTAATTGCGTACAATTTGGTTTCTATCTTAGAGGATCGTTAATGATTTAAAGAATTCGTTTTG
TTTTGCTAAATTCACAACGCTCTTTGCTAAGTTTTAATGAACGTCAAAACTGTTAGTTGTATGCTTGAATAATTCGCTACTAGTTTTAATATCGATTTAAAAATGTACACAT
TTCAAACTCGCAAGTGGAAGAATAATGATTATTAATTTAACTATTTAATGTGTAATAATTACGCGGTTTTCAGGGCTCTTTTTCTTTCGGGGTTAGTTGAAATATCTTTCGA
CTATAATTAAAGCACAGCACAAATGGTTTTGTTGTTTTGTAAGGCTTCTGTTGTGCTACATTTAAGTTTGTTAAAATAGTTCGATAATAGTTCGGTTGTTAAGTATCGGTTT
TGCTATTTAATTTCGCTGCTTGGTGATTGTGATTATCCTTTTGCTTGTAAATTCTGGCTCTTGTTGCTGTTGTTGGTGTTTAAACTCCACGTCCTGCTGTCGTTTTAAGCTA
ATTGTGCGCATCTGAAAGGACAAAGGATAGCAATTGTATGCCTTTAGCGCTTCATAAATAGTTTACAGAATCTTCCATCTTGAGCCAATCAATAAATTATCCCACATACGGC
AGCAGGCGGAAAAAGCCTTTTAGCTGCACAAACGCTTTACA
(SEQ ID NO: 139)

Exon: 3185..2737
Exon: 2223..1965
Exon: 1876..1001
Start ATG: 2869 (Reverse strand: CAT)

Transcript No. : CT41076
ACACACTCACACATCGACACACTCACCTAAGCACACATAGACACTCACAGACACACCGACACACACAACCGAAGGCACAGGACAATTGCATGGGCGTCCGCCATGACCGGAA
GTAAAGTACGGTCACGTCCACGTCCACTGGCCTTTGCATAACAATCAGCCCGCCAAGAAGGTGCGGAGAAAGCGGGTGAAAGCAGCAGCGAAAGCAGGTGAAAGCAGGAGGA
GCCAAGGAGGAGGGAAACGGACAAGGAGCTAAGGGAGCGTATACAACGAGCACAGGAGCGCCCATTTCAACCGACAGTGCCAGCCCGTAAAAATGCAGGAAATGAGCTACCT
ACAGGATAATTCCAAGGTGGAGGCCCTCACCAAGGCGGTGCTCATATCGATACTGGGCGTGGCCATCGTCCTGTCCAACCTCCTCATTATCGCCACCTATGCCAACTTCAAG
GGACCCACAGAGGTGATCAACTACTACCTCTTGTCCCTGGCCATCGCCGATCTGCTCTGCGGACTCCTGGTGGTGCCCTTCTCCGTGTATCCCGCTTTGACCGGAGAATGGA
TGTACGGCGACATCGTGTGTCGCTTCACCGGCTATCTGGAGGTCACTCTGTGGGCGGTATCGGTATACACCTTCATGTGGATATCAGTGGATCGCTATCTTGCGGTGCGGAA
GCCACTCAGATATGAAACGGTGCAGACGAAAACGAGATGCCAATGCTGGATGGTGTTCACTTGGATATCGGCGGCCCTGCTGTGCTGTCCCCCAATCCTGGGCTACTCGATG
CCCATCGAGAACAACATGACGCACATTTGCATGCTGGACTGGGGGAATATGGCAGCATATAGCGCCACCTTGGCGATATTAGTCTTAGGTCCCAGCCTCATTTCAATCGTAC
ACAACTACGGCTATATCTTTGTAATGATGCGTAAGATTAGGTCCGGCGAGCCGATACACGATAAAGAATATGCAACTGCTCTGGCTGAAAATTTATCGAACCCTAGTCATAT
GATGTCATTCGCATTAGTCTTTGCATTCTGGGTGTCCTGGCTGCCATGGATTCTGTTGCGTCTGTACGAGGTGGTCACGGGCGATGTTATCCAAAGTACTCTTATCAACTTC
GCCGTCGTCTGGATCGGCATATTGAATTCGTTTTGGAAAATTCTGATCATGACCAGTATGTCGCCGCAATTCCGTATCGCTTTGAGAGTATTTTGCTTAACCATTTGTTGTA
AGACTAAGGGCCGTTTGCAAGCAGAGCTAATCGGGTTGGACCCAGATGACTAGAGTTAGATTTAGATTTTCATTCTCCTCAAAAAAAAATACAATAATTATCCGTAAAAATC
CCCCTCCCAAACCCATCCTCCGTAATCCCTTTGAGAAGAACTACAAAGTCCTGCAAGTAAATAATTGGTTGTAATTGCATTGAGCCATTTCCTTAACTCTCTATATATCTCT
CTCTCTCAATATATATAAACACTTACTTGCCTGTTTGCCATATAGTTGCATATTTACGGGCCGTTTGGCGCTACGCGACATGCCAATGAAAGCACCACCCCACAGTCACAGC
CACAGAAAAACTGAAA
(SEQ ID NO: 140)

Start ATG: 317 (Reverse strand: CAT)

MQEMSYLQDNSKVEALTKAVLISILGVAIVLSNLLIIATYANFKGPTEVINYYLLSLAIADLLCGLLVVPFSVYPALTGEWMYGDIVCRFTGYLEVTLWAVSVYTFMWISVD
RYLAVRKPLRYETVQTKTRCQCWMVFTWISAALLCCPPILGYSMPIENNMTHICMLDWGNMAAYSATLAILVLGPSLISIVHNYGYIFVMMRKIRSGEPIHDKEYATALAEN
LSNPSHMMSFALVFAFWVSWLPWILLRLYEVVTGDVIQSTLINFAVVWIGILNSFWKILIMTSMSPQFRIALRVFCLTICCKTKGRLQAELIGLDPDD*
(SEQ ID NO: 141)

Name: putative adrenergic receptor
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384685
TAAATATGCATATTGAACGAGGCAATGCAATGCGATTGAGGTCCGGCTGTGCGTTTCAGTTGCGAAATGAGTTGCGAACAGGATTTATATTCAGATTTAGCCATTAGCCAAA
TGCAACACCTTCGACGACGTGTGGACACGTCGACTTTGATTCGGCCAACTGAGTGATCCGGGGTAGTAGTGTCAGATGCCAACATTACTTATGCACTGAAAGAATAAATAA
AGGATTTACATGTGTAGTAATGTATAATTGGGCTGCATGACATGCATGATAATGATCAACTTTCAAGCAATAATAATTCTATGACATCACGACATTCATACGGACAGAAAGA
CGAACATGGATTTACCATTAAAATGTAATTATAATATAATATAATAAAGTATCTAAAGAATTTTTTACTTGGATCTCCTAGCAATTTTAACCATATTCTCAAATGCAAAATG
TGAAGTTCTTGCAAATAGATACACAGACTAATGGGTCCGCATTTGCCAATTAATTGCCTACAACGCCCAGGTTGCTTTGAATCACAGGTGGACATTTGGACAGGTTTGCTCA

GCTCACTGGCAGCTCCACTTTGGCCATTCAAAATTGTTTATCGCTTTGACGAGTTGCAGTAGCAATTTCGACACCGTATGCCCATAATGAAGAAACGGCATTGATAATAACA
AAGTAATGATAAATTAGATGCTGTTGACGTCTAATTCATTTTTGATGATAGTATTGATTTTATACATAAATGACACAATTTCCCTTTTATTTTGTAAAACATTATGGCATTT
TCACTTTAATGGTCCTGTAAAACACAAGTTAGTTGTAAAGCACAGAAAAACTTTCCATTTCGATGACAAGAAGTCCTTTAAAATTCAAATAAATATATAATTCTAATTCTAA
ATAAATAATTCTAAGCAATGGTTTATGTATTTAAACTCCTTAAATTTTGTTCCAGTTTTAACTAGCACATATACATCAAAAGAGCGGAACTAGATAGTCTACAATGATGCA
AGAGACCGGCAATCAAATGGGACAGACCCATATGCACCAGCGGGTTCCTTTCAACGACACGGTTCTGAAGGACTATCATCTCACGAGCACCGATATCGAGAAGTTTGTAAAG
CTATGGCAGGAGTACCAGATGAAGAACATGACCCCTCAGGTGGATGAGTGCCAGGGCTATTGCCAGGGGGAAATATACAACTGGTTGCGAGCCTACAACAGCATCCATGGAT
ACGTTTCGCTGATGGTAAGTAATACTCGTGATTACAACTCCCCCTGGGCACGAAACCCGCAATGTTGGCCAAAACAAAGGAATCGGTTGACTAATGATGAGCAAGGGAAGCC
AAGACGATACACACAGACCTTCGCCGACAGACCACTGATGAATGTAGGGGACTTCTTTACCTTTGAACTCTTGCCAAAGTCGTCCTTTTTTCGAATTAAAGTTATAGCATGC
GCTCAGCTTAACCATCAAAATCACAAATGAAATTAGTGGCCGGGAATCTCGGTCTTTTGTTTATGTCCGGGTTTCCCTTTTCAGGGCTTTTTGAAAAGGATATAAATTATA
TTAGATGTAGTTAAGCCGTTGAGGCGTTAATGTGCTTTTTTCATTATAATTTTCAGTCGTAAGTCATTTTGCGTTGTAGCAGACGGCCTTTACCACAGGAATTGGGTCAGAA
AATACCGAGAATCCATTTCCATCAAGGCTAGTTCATAGCCCATCTTTTCAAATAAATATACAATAAATCAAATACAGTTGAAAAATGTAACTTTATTGCTTTCCTTCGTTCG
CACAAAAAAATAGTTTACCTCTTAAATTGAAATAAAATTCTGAACCAAACTTCGGTTACTGTAGTTCCATTTTTTTTTTTTTTTGCTTGTGTGAGCAATGTTAAGAAATGTAA
ATACGAGAGGTCACGACAGTTGAGTTGCTGCTCACGAGCTGAATAAATAAAATATTCATCCACTTTGGTGTCCAAGGGGAGTGCATTCAGTTGAAAGTATGTGAGCGCTCAC
AGTTTCGATTTGACAAAGGACGCATTTTATGTTGATTCTTTCTTGTTTCTTTAAATCACCTAGATTTGCATTTTCGGGACAATAGCAAACATCTTGAACATAATGGTCCTGA
CCAGAAAGGAAATGGCCAAGACGCCCATAAATAATATACTTAAGTGGCTGGCGGTGGCCGATATGTTTGTGATGCTGGAATATATACCCTATACATCGTATCAGTATATATA
TATGGGGCCAGGTAAGCATACATTTAGTAATAATATTTAATAATAGTCTTGGTTATAGCTTTACTTGTTATTTATTTATTAGTGTTATCAGCTAGATTATTATTATAAAGGT
ATACTCTTGATCAGGATCTACACACGTCTGTCTGTCCGTTTCTAAGGGTATATTTTGTTAACAATTTTGTTAACAATTCCCATATATCCTTTCCCCTTTGCACTTACTTTTA
GTTGAGTAACGAGTATGTGGTAGTCGAGGCAATCGGCTATGACCTTCTCTGTAGGTTTCCTTTCTTTTTAATTTTAAGCCGTTTAAATTTTTTTGAATTATTTAATATTCT
TTGTATGGCTGTTGCAGGTGAAAAAGATTTGAGCTACACCTGGGCGGTTTGTCTGCTGGTTCACATGCACTTCACCCAGATCCTGCATACGATTTCCATCGGATTGACTGTC
ACGCTGGCGGTGTGGCGGTATGTGGCTATCAGGTGAGATGATCCTTCGGCCAGCTATGCATGACTTATTCCGGAATTCGCCGTCATCCATGTAGAGTTTGTCTCCGGTTCAC
ATTCATGATAATCCACATCCGCAATTCGAACAGACATCCGGCTAATCTGATTACCACATCATTGTGGCAATGTGGCTTCTCTTCTCTTCCCCCGTCTTTTGTTTATAGCCAA
ATGTTACCGTCTCTATTCCACAGACATCCGAACGGGGGCTGTGCCAATTTCCTGCTAGCACATTCCCGGGAGGCGATCCTCCTGCCCTTCATCCTGTCGCCGATCCTCTGCC
TGCCCACGTACTTTGTGTTCCAGGTGCGTGAGACGTACGACGTGGACAAGGTCAACTCGGAGGCCATGTATCATGTTTATTTCGACAAGGACTCGGTGCTCTACAGGTGCGC
TTAGATAAGAAGTCTGATGTCATTTATGCATCCTGCGAAGAACCAATTGTAAAATATTTCATCGTGAAAATCCTTAACAGATTCAACTTCTGGATACATTCCGTGCTCATCA
AACTTCTGCCGTGCGGCATATTAATTGTGATCAGTGCAGTACTTATGCACGTTTTGTGCGAGGCATCGAGACGTCGCTTGAAGCTGAGGGATTACAACAATCCCGCCAAATA
TGCCATCCAGCTTAATCTGAATGAAACCAAATCCAAGAAGTAAGTACGCACTTGACTCGTTTAGTGCTTAAGAATTCGCCTCATTTGCAGTTTGCCTTGTTAGTTTCATGAT
CTTTGATTTTTCTGTTTTTGAATGGTCTCGGATTTTGTCAAATCTCTTCTTAGGTTCCTTTTTGCCATCTTTTTTATAAATTTGTCTTTTTGTTTTTGTATCTTAGGCCGCC
TCGCTGTGATCGCCGTAACGATCGGACTACCCTTTTGCTGGTTGCGGTACTGGTTTTGTTTCTGATCACGGAGTTTCCGCAGGGATTACTTGGTCTGCTGTCCGGAGTGATG
GAGAAGTGCTTCTTCGCCCACTGCTATCCGCCGTTTGGAGAGCTGATGGATCTGCTGGCACTGATCAACGCCGCCGTCGGATTCGTGCTGTACGGTCTGATGTCGAAACAGT
TCCGGACCACCTTCCGATCGCTGTTCATGAAGCGGCACTTTGGCAGCACCGAAATGACCAGATTGACCCGCGTAACCACGACTTGCGTCTAGTTGGGCACATCTAAAGATGT
GATCATGGGCAGGAGCAAACACTAGTGAGGGATGAAGTAGGTGAGGGACATCCAAGATGTTCCGGCCCGGCATTGAATTTGAATTTTCTAAAGCATAAACTTTATAAAAGTT
TCGTTGACTAAACAAAGCCTAAATATGAGAAACTGTGAGTGTGGAATTATTGTACGGGCTGACCAGTTGATCAGTTGATTTCCACAATGGTTTACTTTTAGGTGTAGTCCTA
AGGCGTTGACTGGTGCATCGATCGGTTTAGAAATTGATTCTGCTTGTGATCTGGCTGTTTCGCAAAGTGCCATTTCATTTCCAAATAATTACATTGATTAGTTTACCAGTTG
GTGTTTCAAATTCGTATTAGTATATAAGTTTCATAGGCTAAGTCAGCTGTACAAAGCTGTACAAATACGAGACATTTACTTTTCCATAAGATGCAGAAAAAAAAACAGGAAT
CCATCATCAGCAAAACTGTAGCCATGCTAAGAGGACTGAAGAACTCACTGAGCCCCGAGTTGCTCTCTTCTTAGCTGGAATCACGAAAGAATTCTATTTTGATTGTATGG
CGTTAACTGGTTTCAACAAAGTGCCAATGCATTTTAAATTGGCTAATTGAATAACCATATTTTGAGTACTTTTTAAGTCGCTTATACTATTTATAAATAAATGTGAACACT
TTCAAAAAAAAATATGTTTGATGTGAGTTGCATACAAATGAAACGTTTATTGCTTTCCATTACCGAAAAATGAACATCTAATTCTGACTGACAACAGCATTCACCTCCAGTGG
CAATTCGTCCGGCAGAAGTTCCGTGGCCCGATTGTGTCCCAAATCAGACTCGCCCATATCCGCGGAATCCAGCAGGATGCCATTGCCCTGATTCGAGATGCGCAGCAGGGAT
TTCAGCTTGGTGGCTTGGAATTTGTACTGCAGGATAATGAGATTGATAAAATCTGCTCTCGAGGGAGTTTCGACCATAATGTAA
(SEQ ID NO: 142)

Exon: 1001..1246
Exon: 2080..2251
Exon: 2594..2720
Exon: 2936..3130
Exon: 3217..3399
Exon: 3579..3900
Start ATG: 1001

Transcript No. : CT32473
ATGATGCAAGAGACCGGCAATCAAATGGGACAGACCCATATGCACCAGCGGGTTCCTTTCAACGACACGGTTCTGAAGGACTATCATCTCACGAGCACCGATATCGAGAAGT
TTGTAAAGCTATGGCAGGAGTACCAGATGAAGAACATGACCCCTCAGGTGGATGAGTGCCAGGGCTATTGCCAGGGGGAAATATACAACTGGTTGCGAGCCTACAACAGCAT
CCATGGATACGTTTCGCTGATGATTGCATTTTCGGGACAATAGCAAACATCTTGAACATAATGGTCCTGACCAGAAAGGAAATGGCCAAGACGCCCATAAATAATATACTT
AAGTGGCTGGCGGTGGCCGATATGTTTGTGATGCTGGAATATATACCCTATACATCGTATCAGTATATATATATGGGGCCAGGTGAAAAAGATTTGAGCTACACCTGGGCGG
TTTGTCTGCTGGTTCACATGCACTTCACCCAGATCCTGCATACGATTTCCATCGGATTGACTGTCACGCTGGCGGTGTGGCGGTATGTGGCTATCAGACATCCGAACGGGGG
CTGTGCCAATTTCCTGCTAGCACATTCCCGGGAGGCGATCCTCCTGCCCTTCATCCTGTCGCCGATCCTCTGCCTGCCCACGTACTTTGTGTTCCAGGTGCGTGAGACGTAC
GACGTGGACAAGGTCAACTCGGAGGCCATGTATCATGTTTATTTCGACAAGGACTCGGTGCTCTACAGATTCAACTTCTGGATACATTCCGTGCTCATCAAACTTCTGCCGT
GCGGCATATTAATTGTGATCAGTGCAGTACTTATGCACGTTTTGTGCGAGGCATCGAGACGTCGCTTGAAGCTGAGGGATTACAACAATCCCGCCAAATATGCCATCCAGCT
TAATCTGAATGAAACCAAATCCAAGAAGCCGCCTCGCTGTGATCGCCGTAACGATCGGACTACCCTTTTGCTGGTTGCGGTACTGGTTTTGTTTCTGATCACGGAGTTTCCG
CAGGGATTACTTGGTCTGCTGTCCGGAGTGATGGAGAAGTGCTTCTTCGCCCACTGCTATCCGCCGTTTGGAGAGCTGATGGATCTGCTGGCACTGATCAACGCCGCCGTCG
GATTCGTGCTGTACGGTCTGATGTCGAAACAGTTCCGGACCACCTTCCGATCGCTGTTCATGAAGCGGCACTTTGGCAGCACCGAAATGACCAGATTGACCCGCGTAACCAC
GACTTGCGTCTAG
(SEQ ID NO: 143)

Start ATG: 1

MMQETGNQMGQTHMHQRVPFNDTVLKDYHLTSTDIEKFVKLWQEYQMKNMTPQVDECQGYCQGEIYNWLRAYNSIHGYVSLMICIFGTIANILNIMVLTRKEMAKTPINNIL
KWLAVADMFVMLEYIPYTSYQYIYMGPGEKDLSYTWAVCLLVHMHFTQILHTISIGLTVTLAVWRYVAIRHPNGGCANFLLAHSREAILLPFILSPILCLPTYFVFQVRETY
DVDKVNSEAMYHVYFDKDSVLYRFNFWIHSVLIKLLPCGILIVISAVLMHVLCEASRRRLKLRDYNNPAKYAIQLNLNETKSKKPPRCDRRNDRTTLLLVAVLVLFLITEFP
QGLLGLLSGVMEKCFFAHCYPPFGELMDLLALINAAVGFVLYGLMSKQFRTTFRSLFMKRHFGSTEMTRLTRVTTTCV*
(SEQ ID NO: 144)

Name: TRH receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384808
GTTAAATGATTTACAAACAAGCCGACATCTCGGGAACCATTAAAGAGACATTTAAAGTCAAAAAGTGGCCAACTGAAACCCGTTAAAATCCGCTTGCGACCGCAACTTGGCC
CGACTAAGGCAAAAAATAACATAATCCTGTCAAGTCATGCAGACAGGGCGATTTACATTATAATTTTAAATGGCAATGGATTAGTCATCATCAGTACCTGCTTATCTGGTTT
TCCTGATTTTTAAATCAAATCTGTGACTGTATCTCTTACCTTTCCTCTTGTACCTCGTTATTATAAAATGCGAAAACCCTCTGCCAAAGGACACTATAAAACGAATTTTTTA
AAAGACGCCAAACAAAATGGCACTTTTGTTGAGAAGAAAATATGCCGATAAGAACGTTTTACTTGTAAGACGCTGCAGAAAGATTAGGAACGAAAACTATCGACAATGTTT
ACTTGTATAATCTGAGCTTTCTAAAACACAAAAATAGTGTACTTTTACCATACGCTTAATTTCGAAATTAATCCGGGATATACAGTTGGGGATCCAGATTTTAAATTCGTGA
ACACAGCTGCTGGAGTAGGGCATTTCCTGATCGATCTGTTAGTTTTCGGACTTTTACTTTGGTTGTTGTAAACAATATCCGCAAGTTTCTCGCAAAGTGGGTTGCGTTGTGC
CGTGCCTCAGTAGAGCACTACACACTACCTTCGATTGAGTTCGCTGGGATTTGGATTTCGTTCACAGTCCAAGTTTGGACCGTGCTAGGAGAAGATAATGATCCGATCACTC
AATTAACTTCATGCACCCAGTGACGTCATCAGTGCTTTACTATTCGGTGTCTATATAGTCTATTTTGTTTAATTTCTTACACCCTACATTTCGATACGATTCGATTTTCGTT
TGGAAATTGTCTGAAGGTAAACAAGCATAAATTGGACGTTCAAGCGTAAACAAATCGGTTTTCGGAGCAAAATACAAAAATCGTCAGTTGAGAAAGTCTCTCAAATGGCGCA
GTTCTGCATTCTTGGAGTCCTCCTCATTCTTTCTGGAACACACTGCTCGTGGGGATTCCATGAAGAGACACATTATCCGTGTGCCTTTATCGATACGGCAAACATAACGGGG
AGCTACGGCTTAGATGGCCCATTTGTGCACAACTGGACGGTGATTCCCCGTCACTTCGTAGCCGTGTACGACTTTGTCATCGAAAATGGCATCCGCATCCCCGCAAGCAGGC
ATCTCAGAGCCTGTGTCTGTAAGACCAAACCCTGTGTACGGATTTGCTGTTTACGGGGAGAAATCTACGACCTGGAGAAGAGGCAGTGCTTAGTCCCTGTGGCCGGGGTGTC
CAGCCTTCCCTCCCATAGCCATATGGAGGTGGAGTTGGGCAACGGCAGTCTGCGTCTAGTGAAACTGCAACCCCGTTTAGCATCCACGTGGAAACACCATGTGAGCACATG
AAAGCCGTAACCAAGGGCTCGGAATACGTTCACTGGACACTTCATGAGGTCGGTTGTGAAATAAACCTTCCATCACATTAATGGCTTTACATCGCCTTTTCAGAACGGAACC
ATCAGCCACCGAGGACACATATTTTCTAAGCATTACTGCTTCACTCCGCTGCTCCATGGGAACTCCACCTGGGAGTGGCAGCCATTGGCCTGTGCACCAGAAAAACTATATT
TCGTCTTGGGGGTGCGGGAGTGGACATACGCCATATGTAAGTAGCAAGAGGAAACATTTTTCCAGCAACTTTCAGCAATTCATTATCCCAAATATGCACAATACTGTTTTCA
CAACAATACATAAAAGTAATTGCTAAAGCACCAAAAAACAATCGATGATAGCATTAATAGTTTGATAGTCAGAATTTAATTTACCCAGCGTATTCACTTTCCAGGTCTCTTG
ATAGCTATTCTATCAATGTTTATCGTGCTGATGGTCTACCTTATGTGTTCGGAAATGCGCAACAGTTTCTACGGTGTGGCCATAAAGGCGTACGCCATTTGTATGATCTTGG
GATATGCCCTACTGGCTTATCTCACGCTCCACAACCCAGCAAATCTCTCAAATGCGGCTTGTCGTATTCTACGTAAGTGGTAAATTTGAGTATTCTGTGATATTTAAACATT
CCCTTCTTTTTTTTTTTAGCAAGTTTGGCGTTGATGAACTTGGTGCTTTCCTTTTACATTCTGAGCTTTATTGCATTTAAGCTGTACTTGAGCTTTTATGGCGTTGTCTTC
ACAAAACTGATGTTTTGGCTAATATTTACTCCCATCGTACTGGTAGCCGTGGGCTGGTCTTTCTTTGTGGGTTTCAGTTATTACGGCTCTAGGCTCATCTTCGGAGGCGATA
CATGCTGGTTTGATCGTAAGCACTGCCAATCCATATATGTTTTATAATAATTAATTGGTATTCCACATCGCTAGCACGCAATTGGTCCGTTATGATATATTTTTACGCTCCT
GTTTTCGTGGCTTGCGCCATTAGCGGTTTCTTCTATGTCCTGAGCCAGATCTACATCCGTGATCAGCCAGATATCGAGACGGAGAAGAGCTTCGAGTCCATAGAGAAGAACC
GATTCAAATCATTTTGGAAGTACTTTGGCTACACGGCAGTCGTATGGGTCGTTTGCATATGCTCCTTTGCCTTTAACTACTACTGGGAGAACAGATCCCACCTTAACTACGC
TGTAAGCTTCTGCATGGCCTTCCACGGGTTTGCGGCCCTCTATGCGCTGATTGGAAAGAATCAGCAAATCCAAAATTTTCTGCGGCGCATAGATAAGTAAGTTGATAACTAT
GAATTTTAAGTACTACTGTTAATATATACTGTTGCAGTGGCGAAGATACCTGTGAAAACTCAGTTCCGCTCTCGAGTTTTGGTTAACTTTTCTTGTTGTGCCCACTGTAGCG
TTAGTGTAACTGGAAGCAGTTGAGTGTATCGACAATATCTGTGCAATGTTTTTAAAATCACTTATCAGCATAGACAAAAACCATGTATTTTAAATGTAAAATTTTGTAGGAT
GTGTGTAAAATATATTTGTAAAATTTGAACTCGAACAAAAAAGTACACTTCTACCCTTTAAACGAAATTACGATAATAAAGATACAAAAAATAAATATATCTGAAAAAAAAA
TCAGGTCTGGAATTTCTACAATTGTTAAATATGATAACAAGTCCCAACAATCCTTTCAGTATAGAACAAAACATGGAATATACTTAAGCCATGTGAAAGAACACATAACATA
TTAATATTCTAAACAATTTGAGCAGTTTTCTTGAGTTTACGCCTCCAACAATTTATCCTTCACCTTTCTTGAGTTCCAAATTCCTGGACCGTGATCAAGTTTCTCCAGTCTT
ATTTCAAGTCCCGATTTTACTGTACCCTATCCGCATCGCCAAGCCCGGTGTTACTTCCCCGATCTCTCTATATTATAAATGTGCGCCGATTTTATTTAACATCAGAGGTCAA
TGGCCGCAAAACAGACAGAGAGAGAGAGAGAGAAAGTTGGTCAGAGCCACGAGAACTGCATCCCGACAGGGGGCTGTCAAGAGCAAGGAATCAAGTGGAATGTTGGTGGTGC
TCTTTTGAAAAGGCTTCGCAAACTTTCCCCCTGTACACCCCCTTGCCAAGGCCTAATCGCAGAAAACCCACCCGGGTCGGAGTCCCCAATTCATTTTCATGCCGATTATGCT
TCGGGTTTTGGTCCTGCTCTTTTTGGCTGGGCGTACAGCAGCAGGACTCGGGGTACACACTAAGAGAGGTCTAAGAATATAGAGATAGACAAGGAAAGCAGATGGATCGATA
AAACGTAGTAAAAATGCTCGTAGTAATGGAAGCTGGAAATAACATGGCACAAACAGGCCCCAAACTGCATCGACAACACCATCACAATCCAGACGTGCGCACTGGAAAATCC
CGAGAGGTAGGACATGTTTGGCGCGTTTTTCGACTGCATGTGTGCGAGTGTTCTGTTTTGGTTAGATACGTATGAAAAAGTGAAAAAGCCTATATTGACCGCGCCCGGTAGA
TCGCATAGGCAATATCGATATGGGAAATCATTGACTAACATTTGCAGACATTTAGTTTGGGACACATTAAAGTTAATTTGACTC
(SEQ ID NO: 145)

Exon: 1001..1504
Exon: 1560..1716
Exon: 1897..2088
Exon: 2149..2367
Exon: 2427..2784
Exon: 2838..3116
Start ATG: 1001

Transcript No. : CT32762
ATGGCGCAGTTCTGCATTCTTGGAGTCCTCCTCATTCTTTCTGGAACACACTGCTCGTGGGGATTCCATGAAGAGACACATTATCCGTGTGCCTTTATCGATACGGCAAACA
TAACGGGGAGCTACGGCTTAGATGGCCCATTTGTGCACAACTGGACGGTGATTCCCCGTCACTTCGTAGCCGTGTACGACTTTGTCATCGAAAATGGCATCCGCATCCCCGC
AAGCAGGCATCTCAGAGCCTGTGTCTGTAAGACCAAACCCTGTGTACGGATTTGCTGTTTACGGGGAGAAATCTACGACCTGGAGAAGAGGCAGTGCTTAGTCCCTGTGGCC
GGGGTGTCCAGCCTTCCCTCCCATAGCCATATGGAGGTGGAGTTGGGCAACGGCAGTCTGCGTCTAGTGAAACTGCAACCCCGTTTTAGCATCCACGTGGAAACACCATGTG
AGCACATGAAAGCCGTAACCAAGGGCTCGGAATACGTTCACTGGACACTTCATGAGAACGGAACCATCAGCCACCGAGGACACATATTTTCTAAGCATTACTGCTTCACTCC
GCTGCTCCATGGGAACTCCACCTGGGAGTGGCAGCCATTGGCCTGTGCACCAGAAAAACTATATTTCGTCTTGGGGGTGCGGGAGTGGACATACGCCATATGTCTCTTGATA
GCTATTCTATCAATGTTTATCGTGCTGATGGTCTACCTTATGTGTTCGGAAATGCGCAACAGTTTCTACGGTGTGGCCATAAAGGCGTACGCCATTTGTATGATCTTGGGAT
ATGCCCTACTGGCTTATCTCACGCTCCACAACCCAGCAAATCTCTCAAATGCGGCTTGTCGTATTCTACCAAGTTTGGCGTTGATGAACTTGGTGCTTTCCTTTTACATTCT
GAGCTTTATTGCATTTAAGCTGTACTTGAGCTTTTATGGCGTTGTCTTCACAAAACTGATGTTTTGGCTAATATTTACTCCCATCGTACTGGTAGCCGTGGGCTGGTCTTTC
TTTGTGGGTTTCAGTTATTACGGCTCTAGGCTCATCTTCGGAGGCGATACATGCTGGTTTGATCCACGCAATTGGTCCGTTATGATATATTTTTACGCTCCTGTTTTCGTGG
CTTGCGCCATTAGCGGTTTCTTCTATGTCCTGAGCCAGATCTACATCCGTGATCAGCCAGATATCGAGACGGAGAAGAGCTTCGAGTCCATAGAGAAGAACCGATTCAAATC
ATTTTGGAAGTACTTTGGCTACACGGCAGTCGTATGGGTCGTTTGCATATGCTCCTTTGCCTTTAACTACTACTGGGAGAACAGATCCCACCTTAACTACGCTGTAAGCTTC
TGCATGGCCTTCCACGGGTTTGCGGCCCTCTATGCGCTGATTGGAAAGAATCAGCAAATCCAAAATTTTCTGCGGCGCATAGATAATGGCGAAGATACCTGTGAAAACTCAG
TTCCGCTCTCGAGTTTTGGTTAACTTTTCTTGTTGTGCCCACTGTAGCGTTAGTGTAACTGGAAGCAGTTGAGTGTATCGACAATATCTGTGCAATGTTTTTAAAATCACTT
ATCAGCATAGACAAAAACCATGTATTTTAAATGTAAAATTTTGTAGGATGTGTGTAAAATATATTTGTAAAATTTGAACTCGAACAAAAAAGTACACTTCTACCCTTTAAAC
GAAATTACGATAATAAAGATACAAAAAAT
(SEQ ID NO: 146)

Start ATG: 1

MAQFCILGVLLILSGTHCSWGFHEETHYPCAFIDTANITGSYGLDGPFVHNWTVIPRHFVAVYDFVIENGIRIPASRHLRACVCKTKPCVRICCLRGEIYDLEKRQCLVPVA
GVSSLPSHSHMEVELGNGSLRLVKLQPRFSIHVETPCEHMKAVTKGSEYVHWTLHENGTISHRGHIFSKHYCFTPLLHGNSTWEWQPLACAPEKLYFVLGVREWTYAICLLI
AILSMFIVLMVYLMCSEMRNSFYGVAIKAYAICMILGYALLAYLTLHNPANLSNAACRILPSLALMNLVLSFYILSFIAFKLYLSFYGVVFTKLMFWLIFTPIVLVAVGWSF
FVGFSYYGSRLIFGGDTCWFDPRNWSVMIYFYAPVFVACAISGFFYVLSQIYIRDQPDIETEKSFESIEKNRFKSFWKYFGYTAVVWVVCICSFAFNYYWENRSHLNYAVSF
CMAFHGFAALYALIGKNQQIQNFLRRIDNGEDTCENSVPLSSFG*
(SEQ ID NO: 147)

Name: mth-like 8
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384832
AGTCAGTGCCTGGACGGATCGCACGAGTCCTTCTACTAAGCGCTACCAGTGGTGGTCGTAATCCCCCAATCCAAACCCATACCCAAAGCTTCGCTGCATTTTCGAATACAAA
CACTGAACTATCCGTCGTTGTAAAGCTACCTTATCCACTACGATTACTACACTATCCTACAAAAAACTATGCTGGTTAAAAGACTATACACCTAGCTCCATATACTATATAT
ATATATACAATATCGAGGCATGTGTTGTGTGGGCGAGTGAGTGTGTGCGTGTGACTTCTGATTGAGTTGCCATTTGATTCCAAGTATTTTCGAGTGCTTAGTTTTGTAATTT
GTAGCTAGCCGTTAATGTTTGTATAAACTGTTTCATTGGTTTCCCACTCCGCTTTATTCGATAAGCTCGTTTTCTAGATTTTAGCTTTAGCACTCTTAGTTTAATGTTTAAA
GTTAGTTTCGCTGTTAAGTTTACACCGTAGTCGAGGCGTCTTGAGTAGTTCTTTTGTACATTAAGCTTACACTTAGCTCTACATTAGATATACCATGTCGAACCTACGAGTC
TACGAACATACGAACATATATATGTGCGTTGAAAGCGTTGTATTTTTTGAAGCCGAACTTATGCAACCGAAAGGAGACAATAAACAAATGTGCTATAAAGGACGAAAGTGTC
GCAAATAAATGTGAGCCACTGCTAAAGAAGACCGATTTGCGCAAGGTGTTTTACTCGGCGGCGGCGGCAAGAAGCGTTTGGGGCACAAACGACGGATGAGTGATTGCCACTC
GTACCATCATCATGACGATGGCGTGTGGAACGACGTGCGCATAGTAATGACAGCGGATGGCAATCTGAACGATAAGCGGCACCCGATAATGAAACATTAAACTAAACCGTCG
GTAGTCGCTGGCCATAAATCTGTCGGCAGTCTGCTCGCAGAAGCTTTTAGGTGCGGACCCCTTGCGCGGCATACTTTCGGGCTTCCGTTCCGTTCCGTTCGCCAATGGCCCT
GCTCCACTACACCTTCGACCAGCTCGAGTTGTACCTAGAATGGGCGTTCGCCCAGCACGGAGAGGCCACGCCGATTCCCTCGATCCAGCCGTATCCGGGCGTGTTCGTGGGC
GATCTCAGCCAGCTGAACCGCTTCAAGCGACATGCCTTCTCCGCCGTCGTGGGCACGCTCTTCGTCCTCGCCTTCTGCGGCAACCTGAGCACCCTCTATGTGAACAGCCGCC
GGAAGCTGCGGCCCTTCTTCCGGGCCTGCCTCATCTCACTGGCCTGCAGTGACCTTGTCTCCAGCATATTCTGCACCGTATCCTACATGGCCCAGTTCCAGGCGCAGTATCT
GCAGCTCTGGGTGAGGCGGCAGGGATTGAGCAGACTCAACACAAGCTAATTACAAATCTCCTTAGACAATTGGTGGCTTCATGTGCAAGTTTGTGCCGTTCATAACCACCAC
GTCGGTGTTGTCCGGCAGCCTGACCCTGGTGGCCATCGCCTTGGACCGCTACTTGGCCGTGATGCGACCCGTGCTGGGGTTCTGGAGTCCCGACAAACGCTTCAGCACTCTG
AGCATGCTGCTCATCTGGGCATGTTCGATAGGATCCTCTGGCCCGCTGCTGGGCATCTACGACTACCGAAAGATCTATCTCCTGGATGTGGAGGACTCCAGCGAGGAGAGCG
AGGAGGTCGTGACCGCTGTGCCCGAGGAGCTGGTGGTCACCGAGCTGGAAATGGTTCACATGTGCCTGGCCGGGGACGTAAGTCAACGACTGGATTATCGCTTAATACAGTT
GACAGCGATTAAGGGGATTTGGACGGGGGATGTCAAGTTGTTGAATCAATGTTCCGATTATCTTATCAGCCGCGTCGAAAACAATTTATGCAATTCTATAAATATTCCGTT
TTAGCACGACGTCGGACTCTACTACGTCATACTGTTCACCCTGATCTTCCTGCCGTGCATTGTGTCCTTCCTATGGCTGAATGCCGTGATTGCGAGGCAGCTCTGGCTGCGG
AGACACTACCACCAGGAGCAGCAGGAGCAGCATCAGGAGCCCAAGGAGGGTCAGTTTAAGACCATGGCGAATGGCGGGGATCTCTTGATGCCCTCCACCTTGGTCAGTGCCA
TGGGAGTGGCCGTTCCCTTTGCGCTGGACAACACTCCCTTGCCGCCCAAGAGTACCGTCAACGCTCCTGGCAAGAAGACCACTGCTGCTGCTCTAGCCAGGGAGGCCAGGCA
CCGCAAGATGGTGGTGGTGGTGCTGCTAATGATGGCGGTTTTCATATGCCTTCGACTGCCCGCCTGGGTCTTCCTGATCATGCGGCTTTATGGATCCTACTCCGAGCCCATC
GACTGGCTCCTGTACTTCAGCTTCGGCATACTGAACCTTTTCAGTTGCGCCCTAAATCCCATCTTCTACACCTTCCTCACCCAGACCATTAGGACCTTGACGCTGGTGAAAC
ACAAAATTCAGGGATTTCTTGGCTGTCCGCCTGGTAAAGTGCCGGATGGTATGCCCACGGATCAGATGGACAAAAGCGGGTGCTGCTGCGGCCTGAGGCCACCCACCTTCAC
ATGGAGGTGCCATCCCAGCCGCGATAGGGCAGCAGCCACCGTAATCCGGGATGTGGACCAGCCAGATCCGCCCTCCGACCAGGTGCAACCGGATCCGAGTAGCCTACGGCGG
TTCCTCAGCTACAAGCAGGAGGTGTTCACCATTTACAAGCAGTGCGGTGACTCGTCGAGTGCGTCAATCGAGTCCTCCGCCTGATAGGCGGTGGTCATAAATCACCCGGGTG
CATTTGTCGGATACGGGGGCCAACGGTTAGCAGAACCCCAGAACTTGATTGATGGATGGATCCAGCCCCCATTCACTCTGTACAGGCGCAAAAACGGTTTGTCACGGTGGGT
ATATCACCCACGAATACACTAGCACGCATATAAGATATAAGCTCACATATTATGTACCTCATCAAATGTGTAAATATGAGGAATAAACGAATGGGTGTTGGCAGCTAACTAA
CCATTAACACATGGCGCAATCGAAATCGGGAAGGTCTCAGCTATTATTGTGACGAAAACCTAAGACCTTTCAAGTACCACTTAAAACTATATTTAGAGAACTTAAGCAAATT
GTGAGATTAAAATGCAATGCAACTCCTAAAGGTGCGAATATACCTATTCTTAAATATTTTACTCTAGAAATACTGACATCTTTATGTGCTTAGTTTTCGTTCAGACAGACT
ATCTTTGTTATTTCTAAGAACGTAGTAATTAGTTAATGGAATTAATTAGTAACCGAAACTTAAAATTGAGTCATGCCAAAAAAAGAGAATTGGCCTGACAAGACACACATTA
TAGAGTCAGCAGATCGTGGACGAGACAATGAGCCCGCAATGCAATGGAACTAGATTGTATCATATACTAGATACTTCCGCTTGGAACAATGGTTAACAGATCTCGTGAGAGA
TGGGGCTGAAAGCAGTTGTTAATTGTCAAATGTTGAGCACGCCAGTTAAATTTAGCTATGCAGCGAAAGCTCTCTCCCATACTCTCTTCCCATGAACTTTTTAAGCCGCAG
AGCCCGAAGCTCGACTCAAAGCTGAATTCTGCGTTTTCTGGCATTCAGTGTCGGCGGTTTGCGGCGAGCGGTTGGATTCGGATTCTTTGGTCGGTAAAACGCGAAATCTTCT
GTGCGGCGGGTGAAATCTTCGGCAGAAGAAACAACAAAGCGACAAGTGTGGAGTGTGGAGTGCATACCGTACCACG
(SEQ ID NO: 148)

Exon: 1001..1354
Exon: 1410..1757
Exon: 1909..2772
Start ATG: 1001

Transcript No. : CT32957
ATGGCCCTGCTCCACTACACCTTCGACCAGCTCGAGTTGTACCTAGAATGGGCGTTCGCCCAGCACGGAGAGGCCACGCCGATTCCCTCGATCCAGCCGTATCCGGGCGTGT
TCGTGGGCGATCTCAGCCAGCTGAACCGCTTCAAGCGACATGCCTTCTCCGCCGTCGTGGGCACGCTCTTCGTCCTCGCCTTCTGCGGCAACCTGAGCACCCTCTATGTGAA
CAGCCGCCGGAAGCTGCGGCCCTTCTTCCGGGCCTGCCTCATCTCACTGGCCTGCAGTGACCTTGTCTCCAGCATATTCTGCACCGTATCCTACATGGCCCAGTTCCAGGCG
CAGTATCTGCAGCTCTGGACAATTGGTGGCTTCATGTGCAAGTTTGTGCCGTTCATAACCACCACGTCGGTGTTGTCCGGCAGCCTGACCCTGGTGGCCATCGCCTTGGACC
GCTACTTGGCCGTGATGCGACCCGTGCTGGGGTTCTGGAGTCCCGACAAACGCTTCAGCACTCTGAGCATGCTGCTCATCTGGGCATGTTCGATAGGATCCTCTGGCCCGCT
GCTGGGCATCTACGACTACCGAAAGATCTATCTCCTGGATGTGGAGGACTCCAGCGAGGAGAGCGAGGAGGTCGTGACCGCTGTGCCCGAGGAGCTGGTGGTCACCGAGCTG
GAAATGGTTCACATGTGCCTGGCCGGGGACCACGACGTCGGACTCTACTACGTCATACTGTTCACCCTGATCTTCCTGCCGTGCATTGTGTCCTTCCTATGGCTGAATGCCG
TGATTGCGAGGCAGCTCTGGCTGCGGAGACACTACCACCAGGAGCAGCAGGAGCAGCATCAGGAGCCCAAGGAGGGTCAGTTTAAGACCATGGCGAATGGCGGGGATCTCTT
GATGCCCTCCACCTTGGTCAGTGCCATGGGAGTGGCCGTTCCCTTTGCGCTGGACAACACTCCCTTGCCGCCCAAGAGTACCGTCAACGCTCCTGGCAAGAAGACCACTGCT
GCTGCTCTAGCCAGGGAGGCCAGGCACCGCAAGATGGTGGTGGTGGTGCTGCTAATGATGGCGGTTTTCATATGCCTTCGACTGCCCGCCTGGGTCTTCCTGATCATGCGGC
TTTATGGATCCTACTCCGAGCCCATCGACTGGCTCCTGTACTTCAGCTTCGGCATACTGAACCTTTTCAGTTGCGCCCTAAATCCCATCTTCTACACCTTCCTCACCCAGAC
CATTAGGACCTTGACGCTGGTGAAACACAAAATTCAGGGATTTCTTGGCTGTCCGCCTGGTAAAGTGCCGGATGGTATGCCCACGGATCAGATGGACAAAAGCGGGTGCTGC
TGCGGCCTGAGGCCACCCACCTTCACATGGAGGTGCCATCCCAGCCGCGATAGGGCAGCAGCCACCGTAATCCGGGATGTGGACCAGCCAGATCCGCCCTCCGACCAGGTGC
AACCGGATCCGAGTAGCCTACGGCGGTTCCTCAGCTACAAGCAGGAGGTGTTCACCATTTACAAGCAGTGCGGTGACTCGTCGAGTGCGTCAATCGAGTCCTCCGCCTGA
(SEQ ID NO: 149)

Start ATG: 1

MALLHYTFDQLELYLEWAFAQHGEATPIPSIQPYPGVFVGDLSQLNRFKRHAFSAVVGTLFVLAFCGNLSTLYVNSRRKLRPFFRACLISLACSDLVSSIFCTVSYMAQFQA
QYLQLWTIGGFMCKFVPFITTTSVLSGSLTLVAIALDRYLAVMRPVLGFWSPDKRFSTLSMLLIWACSIGSSGPLLGIYDYRKIYLLDVEDSSEESEEVVTAVPEELVVTEL
EMVHMCLAGDHDVGLYYVILFTLIFLPCIVSFLWLNAVIARQLWLRRHYHQEQQEQHQEPKEGQFKTMANGGDLLMPSTLVSAMGVAVPFALDNTPLPPKSTVNAPGKKTTA
AALAREARHRKMVVVVLLMMAVFICLRLPAWVFLIMRLYGSYSEPIDWLLYFSFGILNLFSCALNPIFYTFLTQTIRTLTLVKHKIQGFLGCPPGKVPDGMPTDQMDKSGCC
CGLRPPTFTWRCHPSRDRAAATVIRDVDQPDPPSDQVQPDPSSLRRFLSYKQEVFTIYKQCGDSSSASIESSA*
(SEQ ID NO: 150)

Name: Neuropeptide receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384832
AGGACACATTGGCGTACTTAGTCTAGACCACTGGGTGAGCTTAATGTTTTGCCGTTCGGTTATATATCATGCGGTGATTTGTGGACTTAGCATAAAACCTAACTGAACGGCC
CGGCGACATCCCGACTCTCTGGCGACTTTGGACGGCCACGATAAATCAAGGCGCAGATCCGGGAGCCCGGCATATAGTGGATATAGTATGCGTGGCGAAACTTAATAAGAG
ATGCGACAACAAACAAATGATACAGATAGACAGACGGCAGGCGACAGTCGCGGTGGTGCGAGAGGTGGGTAGGTGGGTAGGTGGGAAATGGGATTGTGGAAAGCGGAAAAAT
CTTTAAAGCAATGCCGCTGCACAGGACATTTTTCACCGCGCCATCCATTCGCAGACGGATCGCCAAAGAGAGACATCAACGTGACTCGCGGGGAGAACAAGTGTTTTGATTT
ATCGCCAGCAATGGCAATGGCAGGGCGGACGGGGATTGCCATCCAATTATACTGCTAACTTTCATTATGTCACTGTGCTATTTCGGGACCGGATTCGCCGCCATATTGGCCC
ATTGTTGCGCCATTGTCTATGCACAGCTCTCTGTTTAGCAGTGACACACTGCGGTTTCGGTGGCAGGACTCATCTCCATCCTGCCGCCTGCCTTAGTCGTCCCTTGCCGCTT
CCTCGGTGGACTCCGTCTGATAAATCGTCAGCGCGGCTTCCCTTTCACATGCTTTGACTTCATTATGGCGTAACAAATCGTCGGCATTTTGCATTCCTCGCTGCGGAAACGG
AAGCGGAAGACATCCCCGAAAGAAACTGCGAACTGGGCCTTTGTGGCCACACGACGAATTGGTCACGCGCCATATAGACAATTCCATCTCAGTTTTCCACATGTTTTGATTG
CTCAACGTTGTCAAGGGACTTTGTGCTGCTGTACTTTTTTCGCTTACACTTTAATAGCTTCCGAAGAGACTCATATCCTTCATATCCTTTCTTGCAGCTCTGCCATGGATCT
AAAGACTGGCTCTCTCCTGGTAGCACCCGCCGTGAAGAACTCCAGCTGCTCGCATCCGCGCTACTCCGGCCACAATTTCATCGCCCACATAGGGATTGCGGAGACCATCGAG
GCGGTGCTCATCCTGGTCCTCACACTGGGCGTCATCGGGGCCAACTGCCTGGTCATCTTCGTCATCAACAACCGCCGCTATGCGGCCTACATACACCAACAGGTGGGAGGCT
GAGTCACTGGCTTAACACCTTGTACATCTTATGTCTCTCCTTGTAGCCCCGTTATCTGCTCACCTCGCTGGCCCTGAATGACCTGACAATTGGTCTGCTAATTACACCATTC
GGTCTAATGCCGGCATTGTTCCACTGCTGGCCATACGGAGAGATCTTCTGTCAGATACAGGTGAGCGGGTCAAGAGTGGGTCAACACCAGGTCAACAAGTGGGGCGGAAGTG
CGATAATTTGTTATGGAACACCCGTCTCCCGTTTGCGGCTGTTGCCACTGCGGTCAATTAACTCTCTCTTAACAATGCACAGTAATTACTGGACATAAATGGCGGTGATGGA
AACGGACAATCCCGATTAACGGGCGGAACAATGTGTAATGCTAGTATTAACGTGGTCGAGAGTGTGTAGGATTCTGCGGTAAGAGTCAGGGTTCCGATGCAGAAGGTGTGAC
AAGCGGAAAGTTATCCCAGCCATATGGTTAGGAAGCTAGTGCACATGGAAAAACCATTGTGCCATTAATTATCATGACGGTGGACAGGGGGCATCGTGGTTAGGGTGTGCAA
CAGAAACTGTTAGTTTATCACTTATCAAGTGTCAAAAAGCATGCAACGAAATCACCTAGAATCCATCACAACAATTGTTGCATACATGCTTAAAAAGGTTGGGTTTTTATGA
AAATCAGAGTTGTAAACATTGTAATAAAATCCAGTTAACACTCATACAACTATTTTCATGGGAACCAGACTCTTGTTATATATTTTATACGATGTCGTTTCACAAAATTTGA
ATTTTCCAACTGCCACTTTAGGGTTATTGGTAAAACATGCACGCATTCCACGCATAACTCTCACAAATATTGAAACACATACTGACGATTAATAATTTTATTAAAAGTTTTT
CGATTCGCCTCGACAAGCGAAAAAGTTTGCCGAAAGCAATAATACAAATAATAAACTTCTAAACGAGCACTAACGCATTTTCTGGCTTTTTCGCGTTTTCCAGGCCCTGCTC
CGAGGGGCGCTTTCCCAGCAGAGCGCCGTCATCCTGGTTTGCATGGCAGTCGATAGATACATGTGTGCCCTGCATCCGCGCCGCTACTACCAGCACTCCAGCAAGAAGGTAA
GATTCCACGTCCTATCGCCTCGCATCGATAGGGAACTTCTATTTTGGGTTGCCCGGATAAGGCACGCTCCTCGGCGATCCTTATCCTGTTGTTCTTGCACTCGCGCTCGCGC
TCGGAAGGACAATAAACAAAGCTCCGGCATAATGACGAGAAATCACGCGCCAATAAAATCGACTATAACAATGGGAAGAGCACCGCACAGCAACTACACAATATGGTCGGCG
AGCCAATTACGCAAGCATAAATAAGGATAATGGGCTTTAAGCAGCTTGACAGTCATTAAACAGGCAAAAAATGGGTGGGCAGCGCCAAAGAGCAAGTGCAAAACAAAGCATC
ATGGAGAATGAAAAAGCAAAGTTGGCAAAAACCAACAAAACCGAAAAAAAAAACAACGGGGGCCAAGTCGAATAGACATTCGAAAAGTGTATAAAATGTGGATAAAAAAAAG
GGAGCAAAGTTGGCCTAGAGATTTTTACGCGCAATGACTTCGGCACAGGGCGACGACCACAGCAAAGTCCTTTTCCAGGACGCAGGCAGGCAAATGAAGCGGCACCTCCCA
CGGAAACCGCAGCCAGGACCAAAGAGCAGCAAACGAGCTCTCAACCAACCCTATTTCCCATTCCCCCCTTCGTAATTTATGGCTCCCGGACTGTAAACAAACGAAAACTAGG
AGGCCAACAGTCGGCTAAGAACACCGAAGTCCGAATGTCCTTGTCCCAAATGAACGGGACTGTTTCAGCTGAAATTGGCGATCAATAAATGTATAGTTTAGAATCCATTTCC
CGCAAAATTCAAGCTAACAAACAGATTTTAGCTTACATATATGATTATTTAACCTTTTGATTTACAGTCTTAATTTCTCGCTGTTTGCTTGGGGATTTTGCAAACAGTTGT
AATTACTAAACTAACTAAAAAAGGAAAATGCCATCAGGCCAAATAAAACAATTGGAAAACGGGAAAGTACCAATAGTGGAAAATAAATTTACATAACAATTTTGTATGCTCC
GAAAATCGGCGCCATTTCGGTTTTTCATGCCATCTGGACGAGTGGGCGAAATTGGCGCACTGCCACGCCCCATTTGTAAAAATGATGGGCGCAGTGTCGAAAATTAATAAA
TGGATGCTGAATGCGAAATACTGGGGATGCTGAAAAACCAATGCTCGGCGGAACTCGGTAATGGTAATTGGGAACTTTTGACTTTCGTTTTTGCTTACTTTCAAAAGTGCCA
TAAATTCAAAGCCAGTCTTGGACATATCTCTTTATTTATAGTGAATTTAACAGCTTGACTCAAAACTCGTCTTGTTGTTCGCCACTCACAATTTGTATTAAATGCACTTGAA
AATTCAATTATTTTGTTATCTAGGATATTTGCTGACTATCCTTACAACTTCAGCAGTCGCCGGGTTGTGCTCTACGAAAAACTCCTGCAATCCTGCAAAAAAAAAAAATAAT
ATTTACATTTCAAAAATTGTACCCAACTGTTTGAGATCGCCCTGCTTTGTTTGTTTCGTTCGGGAAAAGTTTTAAATGTGATAGCGTAAATTAAATTAATCACAGCCATGGT
CCGTTTCAAAGAGACACGGACAAAAGTTGAGCTTCCCTTTTTTCCATGACTTTTTGGCACTTGGGATGGGGCCATTTTCATTTTTTCAGGTGGACAGGTGGACAACATTTTC
CCATTCAGGATGTCAGAATGCTGCATTTAGTAACTCAGCCAGTCAGTGCGATTGTGGCAGCACTTTTCCACTTCCAAAGAGCGGAAGTATCGCTTTCCCTCTGTGGTCTGCA
CCCTTTACCAAGTGATTTTGGCATAATCAAAAGTTTGTAATTCTATTTTTGGCCTCTTGGACTCCCTTGTTGAATGTCAGACGAAATGTGCAGTGAAGCGAGTGGAATTTTA
GCAAATTAGAAAGTTGAAGAGGGAATTTCCGACTTATCTGCACGCCTTAGTGTGATATAAGCCACTTCCGGCTAAATAAACACTTAATATGAATTACCCAACCTTCATTCAT
TTTCAATTCAGTATTGGCCATATTGGCCTGAAATGCAGGGAATACGTGCCTGAAGCCCTTCGCCAGTCATAAATCAGACCGCACCTGGCGGAAATGTGCTCCACTAATTGTG
TATTCATTTTGCAGGGCTGCGTGGCCATACTGAGCTTAACCTGGATTATCAGCCTGACGGTGTTTGGCTTTCTGGTGCTGCCCAAAGGTGAGCTGCCCGAGATTGGCCCGAC
ATCAATGTGCTCTCGAACACGTCAAACCATAAACTATATGCGCAGGTTACTACTTCAACAATACGGGTCTAATGGCCTGCGAGCCCTTCTACAGCAAGCCATCCTACAGGAT
ACTGTCCACGTGCGCCCTCTACTTTCCCACGACGATGGTGCTGATGTACTGCTACGGCTCCAGCTTCCACATGAGCCGCTTCCGGCTAAACGATCCGACCATGCCCCTCACG
GCGGCCGCCCACCACCCCCATCCGCACCCACATCCCACCGCCGCCCAGCAGCTGCAGATGCACCAGCACCAGCAACATCACCAGCAGGCGGGAATGCACTCGCACCTCTACC
ACGGCCACTCGCATCACCCCAGCCACCCCAGTCACCCAAACCACCCCAACCACCACGGCCACCCCCACCATCATGGACCGCCCGTTATGGGCCATCTCAGCATGGCCATGAG
CATGGGACTGGCCGGCATGCCGAACATGACAAACAAGATTACCAAAAAGGTGAGCCAGCTGAACCGGGTGAACTGAGGAAGCCAGGTGGCCTGCTGGCCAAAAGGATGAAGA
TGGATGAAGGAGGATGAAAGAGGATGAAGTGGGCGGGCTCCACAACAGTTTGCTCAGGCGGGCGATTTGGGGATTAAGTTAGCAGTAAATAGAATTTTCTCGGCATCGCCAA
ACAAATTGTTGGGTCTATCAGTATGCACTCGCAGAACTTAAAGAAGTGTTAACAGCTTTTAGTATTTGTGAACGGGCTATTCAATATTGGTACCAGCTAATTCAATTAAATT
CCAGAACTCAGGTAAACTAAACTAAAACTCACTTAAAAAAGAACTCGAGTTGTGCAAGTTACTTTACTTACTTTCCGTGCATAAAAATATTTCCAGAAGCCAATATTACATT
TTAGCAAAATAGAAAGCAGCAGAAGTTGCATTTGAAAATATATAGAATACTCAACTATATTGCTAATAAACAGGAATAGAAATTTCGTTTGGGCAGTATCTACACATAATTG
AATCAACAATAGTCACTTTAATTCAATCGCCGCAAACATTCCCACAGATTGTTCCCATCCAAGAGAGAAGAACTCCAGCGGATCCACCTCGCGCTCGATGGCCGCCATTTCACT
GGGCTTCATAGTGATGGTCACTCCGTGGACGATTCAGGAAATAGTCACCGCCTGCACGGGATCCAAGGCAAGTGTCCTGTACAAAGGATTATTGTGTGACCGCTTTGTGCTC
CGTACAACGCGGCGCATACATAATGAGGTCACGCGTGGCTCATCAGTTTTGATTTTGTGACATTTAATGTTCTATGTCCGTCCACAGCTGCGCCGTTCCTCGATTTCCTGG
TCACCTGGACGGCGCTGAGCAATAGTCTGTGGAACCCGTTCATGTACTGGCTGCTGAACTCCGATTTCCGCCGCATGAGTCGCCAGCTGATGCCGAACAAAGTGAGTATCGG
GTATCCAACAGTGGGAAGATGAATCTCCTCTTCTGGCGCTAATTTACTGCATCTCTGATCTGGAGCGATGTTCACTGTCCGTTTCAGCAGCAAATCTGCCATAAGCCCTGAT
TGATTGATTCGCTTATCCCTATTCCAAACACCACGTGCACACTTTCATTGACCTCCCCGAACGCCAATCCCCAAGCCCAAATATGGGTATTTCGCCCAAGTGCATCCTGCAG
CTCGCAAATTTTCCTCCTGCACAGCGCAAAGTTTGCAAAAACAGGAGCTACATCCTCCCGAAAGCCCCCATAAATCCGTAACCGAAGGTGTAAGCTCCAGAATGCTGGCAGA
CTGACCCGGGCAAAGTAAACAATCAACCGACGCAGCAACAACAATAGCAGGTCCTGCAAAGTGATTTTGAATTATGAAAGAGCAAATCTAAGACGGCATTTACATCCTCGTA
GCCCCACAAACCCCATCCCATAGGTCCTGATTCCATTTTGGTAAACGCAGTCAAGTTGCATGATTTACTTCCGGAGAGCCGGGGTCAGTCTGTGGAATATTTATGGCCAAC
CTTGTGCTCTGCCAGAATAAATGCGCATTTAAAATCCTTATTAATTTTTGTGCGAGTCGGAATTTATGATGCCTCCCTGCCGTTTCAGTGCTTCCCCCACGAGGATACGCCC
GAGCACAAATCAGGCTGTTGTCACATCAATGCTAACGATTTCGAAATCACAACGCTTCCGATTCCCCCGGAGCCGCCATCATCCCGTCCTCCGGGCGGAGCAGGAGGAGGAG
GTGGAGGTCCATCCAGTGGAGGTCCCAGCTCTGGCGGCGGAGTGGCCAGTTCCATTGGCGGCTCCGTCCTTGGAATCTGCGGCCGCTCGCGCACAAACAGCCTCAGCCGCAG
CGCCTCCCAGTACATCCGAGGGACAATGGGCGGAGGAGTCCACAATCACCAGGCCGCGGAAACCTCTCCACCACGAGGCCTGACATTGAGGGCCTCTCCGAGAAGTACTGG
GGCGAGATCCTGGAGAGAACCGTCAGCTCGGGCAACCTGAATGCCATGCAGAAGCACCTGCCGCCCCACCTGCCCTACGCCCACCACCTGGTCCACCAGATCCACCAGCCGC
AGCAGCACCACCAACAGCAGACCACCTCCTTCAGCAAGGCCAGCGATCTGCAGCTGAACCTGAACCTCAGCCAAGCAGCCACGGCTGCCGAGCTGGGCAAGTTCTCCAACTC
GGAGCCGAAGCTGTGCGAGCACCTCTTTCACGACGTAAACTGCGCCAAGAGCAAGGCCGCCAACGGGGATGGGGAGTCCTCCGGCAAGCTGGCCTGCGGTCGCAGCATTCCG
GACATCTAGTACGCCGAGGACGTGATCCTGGCCAAGAACCAACTGCTGGCCAGGCAGGCCAAGTGCTCGCACCATCCGCTGCACCAGCAGGCCAAGACGCGTCTCCATTCCC
ACACGGGCAGCCTGCGCTTCAGCCGGATTCGCCTGCTGCCTGCCACAGTCCGCCGCGGAGAAGTCCGCCTTTCCAATGGCCGAGTATCGCAAGTAGGATCACAGAGGGCACTG
GGATCAAGGATGCCGGGATACGGTACTCCAGGATGTCCAGGATATGGTATAACCATGCACAGAATAGGGAGTGAGAGCGAGCGAGTATCTATAAGATTCCATAGGTGTATGT
GCGTGTGAGTGTCTATATAATACCGCATGATCAAACCATAATATGTGTGTTTTCTATTAGGCTTAAGATCTAGGAGGTCCTTTGGTTGGGTTTTCAATTTAAAGTCTCTGAG
ATTCGCAAAATACCCATGGCCTGTAGAAAGCGCCCTGAAGCCTAACTAAAAGAATATGAAAGATAGCACCTATCCTCAGATATTCCACTCTTTATATAAAGTAATCTGTGTT
GGTTATAACCTGAATGATTGCTGTCCATCTGAGAACTTTCAATACCATTAGCTAGACTTTCTTTACTATACATGGATCTGGGGAATCGGAATTAGCCGAGTTAAGCAGAAGT
TAACCATAGTAAGTGTTCATTCAATGTTTGTGATACCCTAAGTTTCCCAGCCCAACATACATACCTACCAGTAACATATAATATAATCCTAGTTGTTGTGCGTGCGTCGGCT
AAATTAAACGTAGGATATATGGAAACAGAGATATCAGACTTTCCTATCAGAACCCGCTCGACATTCGGAAGCCCTTCGCCAAATTCTCCACCACTTAGGGACTAAGGATCTA
CGTAGACTCTGATATCGGCGGACCTTAAAGCGTTGTGGTTAATCATAGTTGTAGTTCCATAGCTCGAGTCGTATACACCAAATTACTCACTGTTCGACGAATCAAAATCAAA

TCCTAACTTTGTGATTGTCAAGAATTTAAGTCGGTTGATGTGGACATTGCTGTTAAATTAATTTTTATTTGTAGCTAAGTAATTGACTAGTTAAAGCCAAATAAATTTACCG
TTTAAGTAAGTCTAACTTTAGTTAAATAAACGAGCGGTGTGTAACAAATGTATTTCATAGGCAGAGGATTCAAATCGAAATAAATAA
(SEQ ID NO: 151)

Exon: 1001..1222
Exon: 1279..1404
Exon: 2232..2348
Exon: 4495..4567
Exon: 4638..5107
Exon: 5648..5818
Exon: 5912..6037
Exon: 6697..7314
Exon: 7402..7599
Start ATG: 1001

Transcript No. : CT32961
ATGGATCTAAAGACTGGCTCTCTCCTGGTAGCACCCGCCGTGAAGAACTCCAGCTGCTCGCATCCGCGCTACTCCGGCCACAATTTCATCGCCCACATAGGGATTGCGGAGA
CCATCGAGGCGGTGCTCATCCTGGTCCTCACACTGGGCGTCATCGGGGCCAACTGCCTGGTCATCTTCGTCATCAACAACCGCCGCTATGCGGCCTACATACACCAACAGCC
CCGTTATCTGCTCACCTCGCTGGCCCTGAATGACCTGACAATTGGTCTGCTAATTACACCATTCGGTCTAATGCCGGCATTGTTCCACTGCTGGCCATACGGAGAGATCTTC
TGTCAGATACAGGCCCTGCTCCGAGGGGCGCTTTCCCAGCAGAGCGCCGTCATCCTGGTTTGCATGGCAGTCGATAGATACATGTGCCCTGCATCCGCGCCGCTACTACC
AGCACTCCAGCAAGAAGGGCTGCGTGGCCATACTGAGCTTAACCTGGATTATCAGCCTGACGGTGTTTGGCTTTCTGGTGCTGCCCAAAGGTTACTACTTCAACAATACGGG
TCTAATGGCCTGCGAGCCCTTCTACAGCAAGCCATCCTACAGGATACTGTCCACGTGCGCCCTCTACTTTCCCACGACGATGGTGCTGATGTACTGCTACGGCTCCAGCTTC
CACATGAGCCGCTTCCGGCTAAACGATCCGACCATGCCCCTCACGGCGGCCGCCCACCACCCCCATCCGCACCCACATCCCACCGCCGCCCAGCAGCTGCAGATGCACCAGC
ACCAGCAACATCACCAGCAGGCGGGAATGCACTCGCACCTCTACCACGGCCACTCGCATCACCCCAGCCACCCCAGTCACCCAAACCACCCCAACCACCACGGCCACCCCCA
CCATCATGGACCGCCCGTTATGGGCCATCTCAGCATGGCCATGAGCATGGGACTGGCCGGCATGCCGAACATGACAAACAAGATTACCAAAAAGGTGAGCCAGCTGAACCGG
ATTGTTCCCATCCAAGAGAAGAACTCCAGCGGATCCACCTCGCGCTCGATGGCCGCCATTTCACTGGGCTTCATAGTGATGGTCACTCCGTGGACGATTCAGGAAATAGTCA
CCGCCTGCACGGGATCCAAGGCAAGTGTCCTGTACAAAGGATTATTGTGTGACCGCTTTCTGCCGCCGTTCCTCGATTTCCTGGTCACCTGGACGGCGCTGAGCAATAGTCT
GTGGAACCCGTTCATGTACTGGCTGCTGAACTCCGATTTCCGCCGCATGAGTCGCCAGCTGATGCCGAACAAATGCTTCCCCCACGAGGATACGCCCGAGCACAAATCAGGC
TGTTGTCACATCAATGCTAACGATTTCGAAATCACAACGCTTCCGATTCCCCCGGAGCCGCCATCATCCCGTCCTCCGGGCGGAGCAGGAGGAGGAGGTGGAGGTCCATCCA
GTGGAGGTCCCAGCTCTGGCGGCGGAGTGGCCAGTTCCATTGGCGGCTCCGTCCTTGGAATCTGCGGCCGCTCGCGCACAAACAGCCTCAGCCGCAGCGCCTCCCAGTACAT
CCGAGGGACAATGGGCGGAGGAGTCCACAATCACCAGGCCACGGAAACCTTCTCCACCACGAGGCCTGACATTGAGGGCCTCTCCGAGAAGTACTGGGGCGAGATCCTGGAG
AGAACCGTCAGCTCGGGCAACCTGAATGCCATGCAGAAGCACCTGCCGCCCCACCTGCCCTACGCCCACCACCTGGTCCACCAGATCCACCAGCCGCAGCAGCACCACCAAC
AGCAGACCACCTCCTTCAGCAAGGCCAGCGATCTGCAGCTGAACCTGAACCTCAGCCAAGCAGCCACGGCTGCCGAGCTGGGCAAGTTCTCCAACTCGGAGCCGAAGCTGTG
CGAGCACCTCTTTCACGACTACGCCGAGGACGTGATCCTGGCCAAGAACCAACTGCTGGCCAGGCAGGCCAAGTGCTCGCACCATCCGCTGCACCAGCAGGCCAAGACGCGT
CTCCATTCCCACACGGGCAGCCTGCGCTTCAGCCGGATTCGCGCTGCCTGCCACAGTCCGCCGCCGGAGAAGTCCGCCTTTCCAATGGCCGAGTATCGCAAGTAG
(SEQ ID NO: 152)

Start ATG: 1

MDLKTGSLLVAPAVKNSSCSHPRYSGHNFIAHIGIAETIEAVLILVLTLGVIGANCLVIFVINNRRYAAYIHQQPRYLLTSLALNDLTIGLLITPFGLMPALFHCWPYGEIF
CQIQALLRGALSQQSAVILVCMAVDRYMCALHPRRYYQHSSKKGCVAILSLTWIISLTVFGFLVLPKGYYFNNTGLMACEPFYSKPSYRILSTCALYFPTTMVLMYCYGSSF
HMSRFRLNDPTMPLTAAAHHPHPHPHPTAAQQLQMHQHQQHHQQAGMHSHLYHGHSHHPSHPSHPNHPNHHGHPHHHGPPVMGHLSMAMSMGLAGMPNMTNKITKKVSQLNR
IVPIQEKNSSGSTSRSMAAISLGFIVMVTPWTIQEIVTACTGSKASVLYKGLLCDRFLPPFLDFLVTWTALSNSLWNPFMYWLLNSDFRRMSRQLMPNKCFPHEDTPEHKSG
CCHINANDFEITTLPIPPEPPSSRPPGGAGGGGGGPSSGGPSSGGGVASSIGGSVLGICGRSRTNSLSRSASQYIRGTMGGGVHNHQATETFSTTRPDIEGLSEKYWGEILE
RTVSSGNLNAMQKHLPPHLPYAHHLVHQIHQPQQHHQQQTTSFSKASDLQLNLNLSQAATAAELGKFSNSEPKLCEHLFHDYAEDVILAKNQLLARQAKCSHHPLHQQAKTR
LHSHTGSLRFSRIRAACHSPPPEKSAFPMAEYRK*
(SEQ ID NO: 153)

Name: Biogenic amine receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384618
GATAAAGCCCAATTACAGAACCTAATACATACACCTAAAATGACTAATTGACATCCATTGTGCGTATTACATAATAATGAATAATTTAAGTGATAGTTGTTTGAACTTACAG
AACACTTCTTAGTATTGCGGACTCATCCCATCCAACTGTTGACTCATTTCAAATTCTATTTAATTATATTTCTGGTCCTGGCACGAACAATAAATTGAGTTTATATGTATTA
CGCCCCTTCTTGTGGTCTCACTTTATGAAAAAGTTAATTGCGTGTAGTTGTTGGCTTTGCTAGATGGATTTTTATTGTGCCGCTTTACTGTCTGCTTTTTTCTTCATTTTCC
ACTCGCCTTTCATTTTTCATCCGCCGGAAAGCGAATCATTTTTTCGCACGCTGAGCAGTGCAAAGAAAGCAAACACAAGCAAACGAGCAGAATTTCATATAGTTGATTGCAT
TGTCTGCAGGAAGATGTTCGCACGATGATGCAGGACATGCAGGACTAAGGACTTCGAAGCGAATGGTGTGGGCTGGGGCAGGGATAGTCTCATGAAATGCTGATAGCATGCG
CTCAATATTGTCAGTAAATAATACGCGCTTGTTAGAGCCAACTGCATACCAACTAAACTGAACACAGTGAGAATAAACATAACAGAAAGTTGAATTTGCTTAATATATAATT
TCATAAATTTGTCACACCAAAATTCGGATTTAAATTAAAAGTATAATTATATTAATAACATTATTTAATATAATTATAATTATATATAAAAATATATTTTTATATGTATGTT
TCATTTGGTTGCTGAAAACCAGTAACTAACTAATCTTTGTGGGCTATCCAAATCTCTTCATTGCAGAAACCGCAGCGTGCAACATGTGACGTCACAGGCATAAAGTATGGAG
AGCATATTAAAATATTTGCCATTTAGCTGGTTTAGCATGCATTGTTGCATTGACTGCCACATAACAAGTGGAAATTAGCAGGCAGCTACTGTAGACGAATCGAAATGGAAGG
TGGATGGTGGCGAGGAGGAGGAGGAGGAGGACGTTTGGGCGGGAAAGCCATCATGGAAGGACACTCGACACCAAATGGGGCAGCTGCAAGCCACAGAAACAACAGCACAAGG
ACCAATATCGCGACCAACGGCTGTGCCCATTCGGGCATCCTGTTATTTGTGCTGACAGCTATGACACTAACGAGCTTAATAACGCCCACAGAGCAGCTGGCAGTGGCGCCAA
ATGGCACCACACTGCATCAACTGGAGTCCGTGGAGTCCGAGTCGTATCCATCCATAAATGGCACTCAGAATGAAACAATGGTGACCAGTGTGCGACCCCACCTGGACCACAG
GAATCGACCGACGCAGCAGAATGGCAGTCACTATTTGGAGTACGACGATGATGGCCCGGACTGTTCGTACAGCTACAACTTCATCCTGAAGCTCATCACGATGATCCTGTAC
GCACTGGTCTGCATCATTGGACTCTTTGGCAACACCCTGGTGATCTATGTGCTGATGAGCTTCTCCAAGATGCAGACCGTAACCAACATATACATACTGAATCTGGCTATCG
CGGATGAGTGCTTCCTGATCGGCATTCCCTTCCTGCTCTACACAATGCAGGTGGGCAACTGGCCCTTCGGCAACTATATGTGCAAGGCCTACATGGTGAGCACCTCGATCAC
CTCCTTCACCTCCTCGATCTTCCTGTTGATCATGTCGGCGGATCGCTACATAGCCGTTTGCCATCCCATATCCTCGCCTCGCTACCGAACGCCCTTTGTATCCAAGTTGGTT
TCGGCCTTCGCCTGGATGACATCCGTGCTGCTGATGCTGCCGGTTATCCTTTTTGCCAGCACCGTGCAGTCGAGCAACGGCAATGTGTCCTGCAACATCGAGTGGCCAGACA
CTCAGAACTCGCACACCGACTCCACCTTCATTTTGTACTCGCTGGTCTTGGGATTCGCCACTCCACTGACTTTTATCCTGGTGTTCTACTGCCTGGTGATCAGGAAACTTCA
CACCGTGGGACCGAAGCACAAGTCTAAGGAGAAGAAGCGCTCTCACAGGAAGGTCACCAAGTTGGTGCTCACGGTAATAATATATGACATTTCGATTAGTTGTGGAACTAAC
TACTACCTGACCACTCTTGTGTTCAGTTTGATAAATATTATTATTACATTATAAATATATATTTTTTTTGACATTTTGAATTCGACTATAAATTAATTGTTTTAAATTGAT
ATATAGAAAGAATTTTATAATATTTATTCAATACAATTTATGGTTTTTGATTTCTCCACAGGTCATAAGTGCGTACATATTTTGTTGGCTTCCACACTGGATTTCACAGGT
AACTAATTGTATACATTACAGTCTAAAAACATTTATCTTTGTGTACTTTAAAAGGTTTTAAAACCATAAGCTATTTGCCTACTATTATCCTTTTGGATTGATTTTTAAATTA
TTATAGCTAAAACAAATTACGAGATTTGTGGAAATTGACATGAACTGGACTTCTTTGGCAGATCATTTGTAGATTCGCTTTTTATTGACTATTTTCCAGGTGGCTTTGATCA

GCTCCGCTCCTCAACGCTGTGCATCTCGTCTGGAGCTGGCCGTCTTCCTGGCCTGCGGATGCCTCAGCTACTCCAACTCGGCCATGAACCCAATACTGTACGCCTTTTTGAG
CGATAACTTCAAGAAGAGCTTCATGAAGGCCTGCACGTGTGCTGCCCGCAAGGATGTGAATGCCCAGTTGCAGCTGGAGAACAGTTTCTTCCCCAAGTTTGGCAAGGGCAGG
CAATCGGAGCGTCTTCTTGGTGGCAATGGAAAAGGTGGCGCCCAGCGTGGGGCATTAACCAAGAAGAAGTGCTTGGCGACGAGAAACAACAATGCTCCGATGGCCACTACAA
CGACGACGACAACCACCACAACGGGCACGGATGCAGTGACCTGTCTCCAGCCGCCAGTACACCAGGTGCCAGCCGAGATCCAGGTGGGAAATCCGGCCACCGTGCTGGTGGT
CAATGCTGAGACCAACAACTGTAAACCGCCCGTGCTCCACACGGACTTATAAGACCGGGCTCCCTCGATGCCCCTGGAAACGGTGGTCTTCATCGCGCGGAGGTGAGACCAT
CAGGTTGAGCTGGACCTCGATACGGCCATCGATTGCCAGGCAATAGCGAGGCAGCCGGAATGCTTGCTATAGGTCGCCTAAAATGATATTAACATTTTATACCAATATGGCA
GAGCTTTGCAAAAACGGGTAAACTTAATCCCATTTTTATATGATCGAATCTCAGCAAGAAGAAATTTTTAAAAACCCGAATAAA
(SEQ ID NO: 154)

Exon: 1001..2332
Start ATG: 1001

Transcript No. : CT33159
ATGGAAGGTGGATGGTGGCGAGGAGGAGGAGGAGGAGGACGTTTGGGCGGGAAAGCCATCATGGAAGGACACTCGACACCAAATGGGGCAGCTGCAAGCCACAGAAACAACA
GCACAAGGACCAATATCGCGACCAACGGCTGTGCCCATTCGGGCATCCTGTTATTTGTGCTGACAGCTATGACACTAACGAGCTTAATAACGCCCACAGAGCAGCTGGCAGT
GGCGCCAAATGGCACCACACTGCATCAACTGGAGTCCGTGGAGTCCGAGTCGTATCCATCCATAAATGGCACTCAGAATGAAACAATGGTGACCAGTGTGCGACCCCACCTG
GACCACAGGAATCGACCGACGCAGCAGAATGGCAGTCACTATTTGGAGTACGACGATGATGGCCCGGACTGTTCGTACAGCTACAACTTCATCCTGAAGCTCATCACGATGA
TCCTGTACGCACTGGTCTGCATCATTGGACTCTTTGGCAACACCCTGGTGATCTATGTGGTGATGAGGTTCTCCAAGATGCAGACCGTAACCAACATATACATACTGAATCT
GGCTATCGCGGATGAGTGCTTCCTGATCGGCATTCCCTTCCTGCTCTACACAATGCAGGTGGGCAACTGGCCCTTCGGCAACTATATGTGCAAGGCCTACATGGTGAGCACC
TCGATCACCTCCTTCACCTCCTCGATCTTCCTGTTGATCATGTCGGCGGATCGCTACATAGCCGTTTGCCATCCCATATCCTCGCCTCGCTACCGAACGCCCTTTGTATCCA
AGTTGGTTTCGGCCTTCGCCTGGATGACATCCGTGCTGCTGATGCTGCCGGTTATCCTTTTTGCCAGCACCGTGCAGTCGAGCAACGGCAATGTGTCCTGCAACATCGAGTG
GCCAGACACTCAGAACTCGCACACCGACTCCACCTTCATTTTGTACTCGCTGGTCTTGGGATTCGCCACTCCACTGACTTTTATCCTGGTGTTCTACTGCCTGGTGATCAGG
AAACTTCACACCGTGGGACCGAAGCACAAGTCTAAGGAGAAGAAGCGCTCTCACAGGAAGGTCACCAAGTTGGTGCTCACGGTAATAATATATGACATTTCGATTAGTTGTG
GAACTAACTACTACCTGACCACTCTTGTGTTCAGTTTGATAAATATTATTATTACATTATAAATATATATTTTTTTTTGACATTTTGAATTCGACTATAAATTAATTGTTTT
AAATTGATATATAGAAAGAATTTTATAATATTTATTCAATACAATTTATGGTTTTTGATTTCTCCACAGGTCATAAGTGCGTACATATTTTGTTGGCTT
(SEQ ID NO: 155)

Start ATG: 1

MEGGWWRGGGGGGRLGGKAIMEGHSTPNGAAASHRNNSTRTNIATNGCAHSGILLFVLTAMTLTSLITPTEQLAVAPNGTTLHQLESVESESYPSINGTQNETMVTSVRPHL
DHRNRPTQQNGSHYLEYDDDGPDCSYSYNFILKLITMILYALVCIIGLFGNTLVIYVVMRFSKMQTVTNIYILNLAIADECFLIGIPFLLYTMQVGNWPFGNYMCKAYMVST
SITSFTSSIFLLIMSADRYIAVCHPISSPRYRTPFVSKLVSAFAWMTSVLLMLPVILFASTVQSSNGNVSCNIEWPDTQNSHTDSTFILYSLVLGFATPLTFILVFYCLVIR
KLHTVGPKHKSKEKKRSHRKVTKLVLTVIIYDISISCGTNYYLTTLVFSLINIIITL*
(SEQ ID NO: 156)

Name: Somatostatin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013383997
GATAAAAGCATTACATTTTCGAGTCTGCACGATCGTATCATCAGTTGTTTGCAGGACCTTGAGTTCTTCATACAAAGTTTTAGCTGCTCATTTTTAATCTGTTCCTTGTGCA
GAGATTACACTTCTCCGAAAAGAGCCGACACGTAAATTGGAAAAGGGTTAGTGCAACTGTTGTTGCAATTACAATTCATCACATAGCCCGTAGTTCCTCTTCCTGCTGCACA
GAAAGAAATTGGTATGTGGCGATCGTTGCAATGCCATGCAATTCAATGCCCAATTCAAAAATATTTGAAATGTGTAAACAACTCACTTGTTTTCTTTTCTGCGTGCATGTGG
TGCTCCTGGCCATTTGCCATTTCGCCAGGCATAATCCCGACAATCGGTGCGATGTGGGTGATACGAGTGCGAATGCGAATGCGAATGCCAATGATATGTGTATTGTATGTAT
GTGCATACATATGGCGTCGTAGCGTGTGGGCGTGATGTTTGCACTCCGGTGCAGCGGGACTATAAATAAGCCGGGTATTTATTAGGGGCAATCATAAACTGCGTTAACGACG
GGCACTCTGCCAACGAGAATAAGATCCGGTAAACAAAGGCACCGGTCGACGTACGTCTCATTGTGCGATAAGCTTTCAATTAAATCTGACCCGGATAACGGGGCGCATGGGT
GCGGGAAATCGGAAATCGGAAACGAAAACGAAAACGGAAGCGGAAATCGAAATCGAAATGGAAAGAGATCAATTTTCAATTGCAGCCAATGCATGCATGTCCATGGGTCCAA
TGCTAATATCGAAAGACAAAGCTCCGTGTTCTGGAGGTCGTGTTCGCCATGCAGACTCATTGGTTGGTTGGGTTTAACTCGTACAAAGTTTGATGGGTTTGGGATTGTTTTT
TTTTTTTTTTAGAACGATATAACTCGATATAATGCTGGTGCGTAAATAACAACGATGTTCATTGTTCGCTTTGCAGCACATTTACTATGCGGTCGATGGCAAAATGACCCT
CCTGTCGAACATTCTCGACTGCGGAGGCTGTATTTCCGCCCAGCGCTTCACCCGCCTGCTGCGCCAGTCCGGCTCATCAGGACCATCCCCATCTGCACCGACGGCCGGAACA
TTTGAATCAAAATCCATGCTGGAGCCAACATCCTCGCACAGCCTGGCGACCGGACGCGTGCCACTACTGCACGATTTCGATGCCTCGACAACGGAATCGCCGGGAACGTATG
TCCTCGACGGTGTCGCCAGGGTGGCCCAATTGGCCCTGGAGCCCACCGTCATGGACGCACTGCCCGATTCGGACACGGAACAGGTTCTCGGTAATCTCAACAGCAGCGCGCC
CTGGAATCTAACGCTTGCGTCGGCAGCGGCCACCAATTTCGAGAACTGCTCGGCGCTTTTTGTCAATTACACCCTACCACAAACAGGTACCTAAAGCCTAATAACTTTTAAA
TCATATTGCAGTATTTTTTTTTTTTTTTGTTGCCGGGATAGGACTCTACTGCAATTGGACCTGGGACACATTGCTCTGCTGGCCACCCACTCCGGCTGGAGTCCTTGCAC
GGATGAATTGTCCTGGCGGCTTTCATGGCGTAGATACGCGCAGTAAGTAACTGAAATACACTCTAAAAAAAAAGCAGCCGAAAAAGAACATGCTGAGTACAATTAGCATAAT
TAATCGTTGAAGCTGTTTAGCTGTTTAGCAGCCACAAGAAGCTTTATGACTTCTGCCGGCTTATTAGGAGTCAAAAGACATGTCCTACGGTGTTTATCCCATCCACACGGCA
TATCTGGCGGTCACAATCATACGGATCATTAGCTGGGAAACGGTTCAAGTGGTGCACGGCAAACAATTACTTGATTCCGTATTCATCGAAAGGATGTCAATATTTACATGCA
TGACTGGGGGCGTCCAATCAACCGCACTCCCAGTCGGGGCCGCCCCTCCTTGGCTTCCATCGCCGGGGGCTGTCCAGGTGTTTCATGTGCACCCTCGAGCTCCGCCAGTTGA
ACAGGGTTTAAAACACCAGCCAATCGATTGCTTAATTACGATCCTAAATGCTTGCTCTACCCCTCCGATAAATTAAACGTTCTGCAATGGGGAATTGCTATTTAATAACCAT
GGGCGATATACACTGAGCGAAAAGATTCTAAGTAGCTGACTTACCACCCAATATTTTACAAACTTAAAATGCTCTTCCATTATCACTAGTGAATGTATTATCTGCGCTTGTG
GTGCAAAATCTAGGCATATCGAACTATATCTGTTAAGAAATAATTATTCGAGTGCATTGTTGAGCTACGGCCTCATTTCTCCGGGCCGTGTCTAGATTTTTAGCCCCAATAT
TTTCAACTGGACGAGGTCGGTCGCCCATTCGATTGGCATAATTGCACTTGTCAGTCAGTGTGTCAAAGCTCGATTACTGTCATTAATCCAACCCAACCTCCCTCTTTTTCGG
TTCGCAGAATTCGCCATCCGAAAGTGTGAGCTGGATGGTCGATGGGGCAGCAGGCCAAATGCCACGGAGGTGAATCCGCCGGGATGGACGGACTACGGGCCGTGTTACAAGC
CGGAGATTATCCGTCTCATGCAGCAGATGGCAGCAAGGACTTCGATGTGAGTTGAATTCTTTCCACACTACATCATAATACCTGTTTCAGTTATAGTTAAGAGCATTGCTT
AAATGCCAACCAAAATCCCAAATCAAGGGTATGGTTTTGAAATTACCTAATAGGGTTCCAAGTTTTTTAGACACCCTAAAAAGTGATTTCAAAGGCCTGATTACATTTGTTG
ATAGTGGTTTATAGTAGTAGTTTTAATTCCATTTCTTCCAACTGGTTTGCGATATCCCTTAGGCCTACATAGACATTGCCAGGAGGACTCGAACCCTGGAGATCGTGGGCCT
CTGCCTCTCCCTGTTCGCCCTTATAGTTTCCCTGCTGATCTTCTGCACATTTCGCTCGCTGCGAAACAATCGCACCAAGATCCACAAGAATCTTTTCGTCGCCATGGTGCTG
CAGGTGATCATTCGCCTGACCTTGTATCTCGACCAATTCCGGCGGGGAAACAAGGAGGCGGCCACCAACACGAGTCTCTCTGTCATTGAGAACACGGTAGGCTTAATATTCC
CGCATAAGGATAAGGTATAAGGCTTATGTTCGATATTATATGTTATGTTCTACAGCCCTATTTGTGCGAAGCATCCTATGTACTTCTGGAGTACGCTCGTACCGCCATGTTC
ATGTGGATGTTCATCGAGGGCCTTTACCTGCACAACATGGTCACCGTGGCCGTTTTCCAGGGCAGCTTTCCCCTCAAGTTCTTCTCGCGACTCGGCTGGTGTGTGCCCATTC
TGATGACCACCGTGTGGGCGAGATGCACGGTCATGTATATGGACACCTCGCTGGGCGAATGCTTGTGGAACTATAATCTCACGCCCTACTACTGGATCCTCGAGGGGCCACG
ACTAGCGGTCATACTGGTGGGTAATATCAACTGTTGATCGCTTTAATCACCTATTCACCCAACTGAAAACATCCGCATGCAGCTAAACTTCTGTTTCCTGGTGAACATTATC
CGAGTGCTGGTAATGAAGCTGCGTCAATCGCAGGCCAGCGATATAGAACAGACTCGCAAGGCAGTTAGAGCGGCTATAGTCCTACTACCACTTTTGGGTATAACCAATCTCC
TGCACCAGCTGGCTCCTCTGAAAACGGCCACGAACTTCGCGGTCTGGTCGTATGGCACCCACTTTCTCACCTCGTTTCAGGGATTTTTTATAGCGCTAATTTACTGCTTTCT
AAATGGCGAGGTAAGACATCAAGTCACTCAGTTGTCCATTTGCCTAAAACCCTCACTATTTCCAACCAAAATATATATTTTTTAGGTTCGTGCCGTGCTACTAAAGAGTCTG
GCCACCCAGCTGTCGGTGCGAGGTCATCCGGAATGGGCGCCGAAAAGGGCATCTATGTACTCGGGTGCTTATAACACGGCGCCGGATACGGATGCAGTGCAGCCTGCAGGAG

ATCCATCGGCCACTGGAAAGCGGTAAGCATCCGCTTTGTGCACACCCCACCTTGAAACTCCTTACATCAATTTAAAACTCCCAGAATATCACCGCCGAATAAAAGGCTGAAT
GGAAGAAAGCCGAGCAGTGCCAGCATTGTGATGATTCACGAGCCTCAACAGCGCCAGCGACTGATGCCCCGGCTGCAAAACAAGGCGCGGGAAAAGGGCAAGGACCGGGTGG
AGAAGACGGATGCGGAAGCGGAGCCGGATCCGACCATCTCCCACATTCACAGCAAGGAGGCGGGCAGCGCGAGATCGCGAACTCGCGGCTCCAAGTGGATAATGGGCATCTG
CTTCCGGGGTCAAAAGGTACTAAGAGTACCGTCAGCGTCATCCGTGCCACCCGAGTCAGTTGTATTTGAGTTGTCAGAGCAGTAGACCAAAACCTAGCCTAACTAATAGATG
TAGCCATAGCCGTAGCCAAATAGCCGTAGAAGTAGCCAACGCCATGTAGCCCACCACCACAGTAGCCAGCACCCACACAATCGGACCGGAAATGATCCCAAAGATAACCCAG
AACCCGACTATTCCTGCATCTCCCATCGACCCACCAAGCAAGCCAGCTGGAAACAGCACATATGTAAAATACCCCAACAAAATATCGCAGTTCTTATACCCAAGATCATTTT
GAACAGAAAGTAAGCTCATTGTTTTGTAATTCAAATTAAACCAATAGACATTTAATTTAAATAGTTATTTTCGCTTGCCTTCGGTTCTTGCTCCACTGTACTTATGTAAATA
CATACATACCTTTATTTGTGTAAGCCATATCCAAACCCTTTCCAAAATCTAGAAATCTCCTATCCTTTCGAAATCGGCTAATCCTACGAAAATTCAAATTTCCCTACAAGTT
TTACAAAATATTTTTGATGACAATAAATTTCCAATAGCGTCACCTTCTGTTCTGGACACGCAAATACATTCAAAAACACTCGCACACGGTTCTTACTCGATGAATTTTTCAG
CAAAACTTCAGCGGCATGTTCTTTTACTTCTACTATTCATATTTAAATATTAGATTGAAATATTTAAGTGAATTAACACAAACTTGCGGTAACTGCGCAGTGTGGCCAGGCT
TTGTGAGCCATTAGAACTTATTACATAACATAATTGAAAAGTGACATTGTTTAAAGGAACTCACCTTTAATCAACTGTTGAACTTGCTGTTCATCGGATTCCGTTATTAAGT
TTTAAGATTTGGGTTTACTTGTATTTACATTTTCTTTACCACTTTGATTAGCTTACACTTTCCGTTCCTGAGCTGGTCACACTGCCAGCTACAATCGATCACGTTCACTGCA
ACTGCCTTGAATTTTGCTAGCTGCGCTTTGTTGTGCAACAAAAACTTTGCCACAATCGATATTTAACTAATGGGGTC
(SEQ ID NO: 157)

Exon: 1001..1322
Exon: 1500..1610
Exon: 2472..2623
Exon: 2863..3120
Exon: 3192..3488
Exon: 3555..3818
Exon: 3894..4054
Exon: 4117..4453
Start ATG: 1001

Transcript No. : CT33238
ATGACCCTCCTGTCGAACATTCTCGACTGCGGAGGCTGTATTTCCGCCCAGCGCTTCACCCGCCTGCTGCGCCAGTCCGGCTCATCAGGACCATCCCCATCTGCACCGACGG
CCGGAACATTTGAATCAAAATCCATGCTGGAGCCAACATCCTCGCACAGCCTGGCGACCGGACGCGTGCCACTACTGCACGATTTCGATGCCTCGACAACGGAATCGCCGGG
AACGTATGTCCTCGACGGTGTCGCCAGGGTGGCCCAATTGGCCCTGGAGCCCACCGTCATGGACGCACTGCCCGATTCGGACACGGAACAGGTTCTCGGACTCTACTGCAAT
TGGACCTGGGACACATTGCTCTGCTGGCCACCCACTCCGGCTGGAGTCCTTGCACGGATGAATTGTCCTGGCGGCTTTCATGGCGTAGATACGCGCAAATTCGCCATCCGAA
AGTGTGAGCTGGATGGTCGATGGGGCAGCAGGCCAAATGCCACGGAGGTGAATCCGCCGGGATGGACGGACTACGGGCCGTGTTACAAGCCGGAGATTATCCGTCTCATGCA
GCAGATGGGCAGCAAGGACTTCGATGCCTACATAGACATTGCCAGGAGGACTCGAACCCTGGAGATCGTGGGCCTCTGCCTCTCCCTGTTCGCCCTTATAGTTTCCCTGCTG
ATCTTCTGCACATTTCGCTCGCTGCGAAACAATCGCACCAAGATCCACAAGAATCTTTTCGTCGCCATGGTGCTGCAGGTGATCATTCGCCTGACCTTGTATCTCGACCAAT
TCCGGCGGGGAAACAAGGAGGCGGCCACCAACACGAGTCTCTCTGTCATTGAGAACACGCCCTATTTGTGCGAAGCATCCTATGTACTTCTGGAGTACGCTCGTACCGCCAT
GTTCATGTGGATGTTCATCGAGGGCCTTTACCTGCACAACATGGTCACCGTGGCCGTTTTCCAGGGCAGCTTTCCCCTCAAGTTCTTCTCGCGACTCGGCTGGTGTGTGCCC
ATTCTGATGACCACCGTGTGGGCGAGATGCACGGTCATGTATATGGACACCTCGCTGGGCGAATGCTTGTGGAACTATAATCTCACGCCCTACTACTGGATCCTCGAGGGGC
CACGACTAGCGGTCATACTGCTAAACTTCTGTTTCCTGGTGAACATTATCCGAGTGCTGGTAATGAAGCTGCGTCAATCGCAGGCCAGCGATATAGAACAGACTCGCAAGGC
AGTTAGAGCGGCTATAGTCCTACTACCACTTTTGGGTATAACCAATCTCCTGCACCAGCTGGCTCCTCTGAAAACGGCCACGAACTTCGCGGTCTGGTCGTATGGCACCCAC
TTTCTCACCTCGTTTCAGGGATTTTTTATAGCGCTAATTTACTGCTTTCTAAATGGCGAGGTTCGTGCCGTGCTACTAAAGAGTCTGGCCACCCAGCTGTCGGTGCGAGGTC
ATCCGGAATGGGCGCCGAAAAGGGCATCTATGTACTCGGGTGCTTATAACACGGCGCCGGATACGGATGCAGTGCAGCCTGCAGGAGATCCATCGGCCACTGGAAAGCGAAT
ATCACCGCCGAATAAAAGGCTGAATGGAAGAAAGCCGAGCAGTGCCAGCATTGTGATGATTCACGAGCCTCAACAGCGCCAGCGACTGATGCCCCGGCTGCAAAACAAGGCG
CGGGAAAAGGGCAAGGACCGGGTGGAGAAGACGGATGCGGAAGCGGAGCCGGATCCGACCATCTCCCACATTCACAGCAAGGAGGCGGGCAGCGCGAGATCGCGAACTCGCG
GCTCCAAGTGGATAATGGGCATCTGCTTCCGGGGTCAAAAGGTACTAAGAGTACCGTCAGCGTCATCCGTGCCACCCGAGTCAGTTGTATTTGAGTTGTCAGAGCAGTAG
(SEQ ID NO: 158)

Start ATG: 1

MTLLSNILDCGGCISAQRFTRLLRQSGSSGPSPSAPTAGTFESKSMLEPTSSHSLATGRVPLLHDFDASTTESPGTYVLDGVARVAQLALEPTVMDALPDSDTEQVLGLYCN
WTWDTLLCWPPTPAGVLARMNCPGGFHGVDTRKFAIRKCELDGRWGSRPNATEVNPPGWTDYGPCYKPEIIRLMQQMGSKDFDAYIDIARRTRTLEIVGLCLSLFALIVSLL
IFCTFRSLRNNRTKIHKNLFVAMVLQVIIRLTLYLDQFRRGNKEAATNTSLSVIENTPYLCEASYVLLEYARTAMFMWMFIEGLYLHNMVTVAVFQGSFPLKFFSRLGWCVP
ILMTTVWARCTVMYMDTSLGECLWNYNLTPYYWILEGPRLAVILLNFCFLVNIIRVLVMKLRQSQASDIEQTRKAVRAAIVLLPLLGITNLLHQLAPLKTATNFAVWSYGTH
FLTSFQGFFIALIYCFLNGEVRAVLLKSLATQLSVRGHPEWAPKRASMYSGAYNTAPDTDAVQPAGDPSATGKRISPPNKRLNGRKPSSASIVMIHEPQQRQRLMPRLQNKA
REKGKDRVEKTDAEAEPDPTISHIHSKEAGSARSRTRGSKWIMGICFRGQKVLRVPSASSVPPESVVFELSEQ*
(SEQ ID NO: 159)

Name: Calcitonin receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013385192
GCTTACAAAACCAAAAACCAGATTTTGTGCTTCAAATCACAAACATAACGAAAGAAATATGAATGGAAAACAGAGGATTTTATCATCACTTGAAGGCTGACAAAACATTTGG
CACAATTTACATTGTCCGGCGATTTATTGCCGTCTAGAGACAAAAAACACTCCGGAAAATTGTCACTTTGCTTGCAGTAAGAAAGTAACTTTTAATTAGAGGTTACTGCAGC
TATGCAGCTGCTACTGCGGTTGCTACTAAAAGCAAAAGTGAAGATATATCCTTTGAATCTCAGCACAAATTGCATTCTCCGTATTTTCTCATTTGTCTAATGGAGTTATGTC
ACCAAGAAATGCGTGAGTGTTACACTCACAAAATATGTTCTATTCCCAAGAAATTCTTCTCTGTAAGTAAGTGTAAAAGTGTGTACACTAACTCTGTTTGTATTAAAGACAA
ATGCATAAGAATGCCTTAAATTCGTCTACGTGTAACATGTAAGCCAGAATTTTTGTTGCATGTTACACGTTTTCCCACAAATTGCCAAAATCATCTGCTTCTGCCGGCCCAC
TAGTCGTATGCGTATTATTGATTACGGAACGAGGCCTTGTTTTCTTACCATTTCGGACAAATGGTAATGACATTTCATTGTTGATTCGTTTGAAATCGTACCTTTGACCTGC
CCTGAACTTGGACATCGTTCTTTGCATGTGGCATGCAATGCCATTGTATTTGGGGAATTTCGAAGCTTAAATTCTGGTAAAATTCAAACCACAACACTGAATTCGTTTTTA
TTCACCTAACGCCTGATGTTTATTTGCAGGTAGCAGGTAGCACTTTCGCCACGCAGCCTGCATATTTGAGTGTTGGAATTTCCGAAAAGGGAATTTAAATGAAACGCCAGAG
GAGCAGCGGGGCGCACTGAGTTGGCAGCCGGACACCGCGGTCAATCAAACGGAGTCCTTGCAGGACATCCCACCACCGACTACCATCTGGCCTTCATTGACACCATGGTCAC
GAACATGTCGCAGCCGCATTATTGCGGCACCGGCATCGATGATTTCCACACAAAGTGAGTAGAACGAGGACGTGGACTTAGGCCTGGGCACATAGCCCCTGGGTATGTCTGC
ATATTAATTAGTTTTTTAAGTGTAATTTACATAATGATTTGCCAGCCGGGGGGCGCATAATGAAATCATGTGTGTAATTGCCTACGGCCGTGGTGGCAAAAAGGCTAACGAA
GTGTGCCTAATGTGGATGGGCCCCTTTCGGATTAAACTGAACAAATACTCGCATAAATCCATAAAGCAGATGTTAAAGGTCGGAGATGAACATTTTGAAAAATATATTGAGT
TCAATCTAACCAACCATGTTTGTTTATTCGGTTTCTTTCTCTACATAAGCTCTAATTACTTTAAGAATTTAATCATACCCAATTTTGTGCTCAAGTGAAACTGGTTTTAAATA
TTTATCCCAATTTTTTTTGTCTTTCTCCCCAACAGCTACAAATACTTTCACGGTTACTTCTCGCTGATTGTCTGTATCCTGGGAACCATTGCGAATACCCTAAATATCATAG
TGCTGACCCGACGGGAGATGCGCTCCCCCACAAATGCCATACTCACGGGTCTGGCTGTCGCCGATCTGGCTGTGATGCTGGAATACATACCCTATACGGTGCACGACTATAT
CCTCAGTGTAAGGCTGCCGCGAGAGGAGCAGCTCAGCTACAGCTGGGCGTGCTTCATCAAGTTTCACTCGGTATTTCCCCAGGTGCTGCACACCATCTCCATTTGGCTAACG

GTGACGCTGGCAGTTTGGCGGTACATAGCGGTAAGCTATCCGCAAAGGAATCGCATCTGGTGTGGAATGCGAACTACTCTGATCACCATAGCCACGGCCTATGTTGTCTGTG
TCCTGGTAGTGTCACCTTGGCTGTACCTAGTCACAGCCATTGCCAAGTTCCTAGAGACTTTGGATGCCAATGGCAAGACGATTGCCTCAGTGCCATTGAGTCAATACATTCT
GGACTACAATCGGCAGGATGAGGTGACCATGCAGGTCATGTCGAGTACAACGCCAGATGTTTCCTGGGCGATACCAAGTGATTCGGCCAATGGAACTGCAGTTAGCTTGCTA
AGTCTAACCACAGTGATACCCCTAACCACATTAAGCACTGGAGTAACCACATCCTCGTCGTTGGGTGAGCGCAATGTGACTGTCTATAAGCTGTATCACAGCGCACTGGCGC
TGCGTGATCGGCAGTTCAGGAATGCGACCTTCCTTATATACAGTGTCCTGATCAAGCTGATACCCTGCTTCGCACTGACCATTCTGTCTGTGCGGCTCATCGGTGCTCTGTT
GGAGGCCAAAAGGAGGAGGAAGATCCTGGCCTGTCATGCAGCCAACGATATGCAGCCAATTGTCAATGGAAAGGTGGTGATTCCGACGCAACCCAAGAGCTGTAAACTGCTG
GAGAAGGAGAAGCAGACCGATCGCACCACGAGGATGCTTCTGGCGGTACTGCTGCTCTTCCTGGTCACCGAGTTTCCACAGGGCATTATGGGTCTGCTGAATGTGCTCCTGG
GCGATGCCTTCTTTCTGCAATGTTACCTAAAGCTGAGTAAGTATTTTAGGATTAAGACACCCACTCCTCCGCTGTCTGTAACCCGTAATTGCTTTATTGGCTATGATAATTG
ACTGAATTTATGGTTGGTGTTTAGGTGGGCCCAGTTTGGCACATTCATCAATCATACGCACCGTTGGCTGCCATAATTGCACCACAAATTTGAGGCTTCAATGAAGGCCTGG
CATTGACGTGTTCAATAAGTGTACATTTTGTTTATCTCGTACACCCAAAATAAAGTAGTCTTACATTTAAAACTTCGCAGATACTATTAAATTGAGAGATTTTTTTCTTTAA
TTTGAGTAAAAGTATAAAATTGAAAAAAAAATATTTTAAACATTTAATATGTAGTTTTTAGCTAGCATATGTTTTTTCAAGTGTAGGAGGTATAAAATCTAATCGCCCTTGA
ACTTTATGAAAGCCCTGCCCATAGATGGCGTGTTGGTGGAACCCGACATTTATGGCATTTCGATTTTTTCGTTTCAGGTGACCTTATGGACATCTTGGCGCTTATTAATTCG
AGCATCAACTTCATCCTGTACTGTTCGATGAGCCGCCAGTTCCGGAGCACGTTCGCGCTCCTCTTCCGTCCGCGCTGGCTGGACAAATGGCTGCCGCTGTCGCAGCACGACG
GCGAAGGGAGGGTGGGCGGAAGTGGCGGCCTGGGCGGCTACGGCGGATATGGACGGCAGCGGTTGCTGCACACGGATGCCGTTAGCAAGAGCATGGCCATCGATCTCGGGCT
GACGACCCAAGTGACAAATGTGTAGCAGGAGAGCAGCGGCCGGGCGGCGATGTCAGCCGCAGCCGGCGGAGCAGCTGCATCGGTGGCTCTGGCCCTGGCAGCCACTGATGTT
GATGGATGTCCGCCTGCCACCGATGCTGCTGTTTCCACTAACGACATCAGCCTGGTCGAGAAGCTACATTTGCAGCCCAGTCCAAGGGGGACTGCGATATCCAGTGGCCAGC
ATCGAAGGCGGCGCAGTGGGAGTGGCACCAAGTGCATCTGGCCCACGACGGACTGGCTGAGAAAGCTGCGTAACCAGAAGGCTAGAGAAACGGAGCAATCCTCCGAACAGGA
CATAGAGCTGGGCAAGAGCTCCATCAACAGGCGCAGCAGTGTTCTTCTAATGGTATTACTAAGCAGCTCCGATGAGGTTAAAGCCAAGGCGGTTTTGGTCAGTGAACAGCCG
CCGAGTCCAGCGGACGAGGATGTGGAAGACGCCATTGACGCCCTCTGGCTGTGAGACGACCCCTCACTTCCATTTTGTACGCCCTGTAAATTATCCAAGTACCCACTAATAT
TTAACTACCGTGTACAGACCAATTATTGAGACAAATTGTTATGCGCTGATTCCATTATTTTCCCGCCAGTGTTGTCTTGGCGAATCGAGAGCCCCTTTCGCTGAGACTTTGC
TTTGCTGGAAAATGAGTGGATTTTCCTTGAAAGGACTCCCTGGAAGCAGCCTGCAACATTTTGTGACCACATTTGATAAGTGATATACTTGACTAAGTTTCCCCCCTTAATA
GTTGTGTAAGTGCTAAGCGCGTCTCTAGATATAAACGAATTAGTCATAGTCTAACTATTAATTAAGATATATTGATGAAATAATTCGGAAGCTAAGAGATTAAATTTCTTTT
TATATTGAATAAATGTTTTTTCGAAGTAAACTTGATTCAATTCACTTATTGTAAGTGTTATTGCGACACTTTGTTTTAATTAAACGGAATTTATATTGAACTTTGATAAACG
AATGCTTTTGAATGCAT
(SEQ ID NO: 160)

Exon: 1001..1062
Exon: 1492..2093
Exon: 2172..2612
Exon: 3102..3385
Start ATG: 1001

Transcript No. : CT33298
ATGGTCACGAACATGTCGCAGCCGCATTATTGCGGCACCGGCATCGATGATTTCCACACAAACTACAAATACTTTCACGGTTACTTCTCGCTGATTGTCTGTATCCTGGGAA
CCATTGCGAATACCCTAAATATCATAGTGCTGACCCGACGGGAGATGCGCTCCCCCACAAATGCCATACTCACGGGTCTGGCTGTCGCCGATCTGGCTGTGATGCTGGAATA
CATACCCTATACGGTGCACGACTATATCCTCAGTGTAAGGCTGCCGCGAGAGGAGCAGCTCAGCTACAGCTGGGCGTGCTTCATCAAGTTTCACTCGGTATTTCCCCAGGTG
CTGCACACCATCTCCATTTGGCTAACGGTGACGCTGGCAGTTTGGCGGTACATAGCGGTAAGCTATCCGCAAAGGAATCGCATCTGGTGTGGAATGCGAACTACTCTGATCA
CCATAGCCACGGCCTATGTTGTCTGTGTCCTGGTAGTGTCACCTTGGCTGTACCTAGTCACAGCCATTGCCAAGTTCCTAGAGACTTTGGATGCCAATGGCAAGACGATTGC
CTCAGTGCCATTGAGTCAATACATTCTGGACTACAATCGGCAGGATGAGGTGACCATGCAGGTCATGTCGAGTACAACGCCAGATGTTTCCTGGGCGATACCAATAACCACA
TCCTCGTCGTTGGGTGAGCGCAATGTGACTGTCTATAAGCTGTATCACAGCGCACTGGCGCTGCGTGATCGGCAGTTCAGGAATGCGACCTTCCTTATATACAGTGTCCTGA
TCAAGCTGATACCCTGCTTCGCACTGACCATTCTGTCTGTGCGGCTCATCGGTGCTCTGTTGGAGGCCAAAAGGAGGAGGAAGATCCTGGCCTGTCATGCAGCCAACGATAT
GCAGCCAATTGTCAATGGAAAGGTGGTGATTCCGACGCAACCCAAGAGCTGTAAACTGCTGGAGAAGGAGAAGCAGACCGATCGCACCACGAGGATGCTTCTGGCGGTACTG
CTGCTCTTCCTGGTCACCGAGTTTCCACAGGGCATTATGGGTCTGCTGAATGTGCTCCTGGGCGATGCCTTCTTTCTGCAATGTTACCTAAAGCTGAGTGACCTTATGGACA
TCTTGGCGCTTATTAATTCGAGCATCAACTTCATCCTGTACTGTTCGATGAGCCGCCAGTTCCGGAGCACGTTCGCGCTCCTCTTCCGTCCGCGCTGGCTGGACAAATGGCT
GCCGCTGTCGCAGCACGACGGCGAAGGGAGGGTGGGCGGAAGTGGCGGCCTGGGCGGCTACGGCGGATATGGACGGCAGCGGTTGCTGCACACGGATGCCGTTAGCAAGAGC
ATGGCCATCGATCTCGGGCTGACGACCCAAGTGACAAATGTGTAG
(SEQ ID NO: 161)

Start ATG: 1

MVTNMSQPHYCGTGIDDFHTNYKYFHGYFSLIVCILGTIANTLNIIVLTRREMRSPTNAILTGLAVADLAVMLEYIPYTVHDYILSVRLPREEQLSYSWACFIKFHSVFPQV
LHTISIWLTVTLAVWRYIAVSYPQRNRIWCGMRTTLITIATAYVVCVLVVSPWLYLVTAIAKFLETLDANGKTIASVPLSQYILDYNRQDEVTMQVMSSTTPDVSWAIPITT
SSSLGERNVTVYKLYHSALALRDRQFRNATFLIYSVLIKLIPCFALTILSVRLIGALLEAKRRRKILACHAANDMQPIVNGKVVIPTQPKSCKLLEKEKQTDRTTRMLLAVL
LLFLVTEFPQGIMGLLNVLLGDAFFLQCYLKLSDLMDILALINSSINFILYCSMSRQFRSTFALLFRPRWLDKWLPLSQHDGEGRVGGSGGLGGYGGYGRQRLLHTDAVSKS
MAIDLGLTTQVTNV*
(SEQ ID NO: 162)

Name: TRH receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384809
ATAATTCTTATGAAAAATATAAAAATGGGCTGATTAAAATGATTGGTATTATGGAAATTATTATTGAAAAAATATGAAATTTACATTTTAAAAAGTGTTATTCGCGTTATTA
CTCAGTTTGTAAATAGCGGGCATCAATTGTGCATAGTTTGAAGGTAAGTTGCCCAATAAATGGCGCAAAGCTCGTAACCGCGAATTATATATGTATGTACATATACGATAAA
AATAAAGTGTGGCCTTTGCTTGGATAAAAAAATCGTCCGAGTCGCCGGATATCCATATTATCGCTTTCAGAAAATTCGATATCATCGATAGGTGTTTCCCTTTCAAAAGTCT
TCGTCCGATTTGGTCCGTTTGACTACTTTTATCTGGTGCATACCCGATTTGACTGCTTGGACGGAGTGGGATGTGTGAATGAACACTAAAACGCTTGAAACTTGAGAAATTA
GGGAAGTAGACGGGTAGCGTCAGTTTGTTGAGAAGATTAACAAGCCCACAAACTGCACAAAACTTGCAAGGCCATTTTATTATTTGTCTATCCATTTTCGATCGATATGGC
AACAACATTTAAGCCATGCCCACTCTAATGCGTCAATAAATGCGACTAATCCAGGCGGGCCCCCAAGTTAACTTGGCCCTCCTCATCGAGTCAAGCGCTTCCTTCGAAGAAT
GATTTACCGTTGATCATACCGGGAAGCCTTTTTTGGTCTTCAGGGGAGGTAGAGACTACATGTGGTATTAAGATGAAAAAACTAAATGCGTACATCAAATCAAAAGTATTCT
TATTAATCCAAGTGAATTATAATATTTATTTTGGCACCTTGTAAATGTGACAAGTGATTAGAAATGTGGTTCTTCTTAAAGATATAAATGTACATCATAATATGCTAGAGTA
CTGAGGCACTATTATTTATATACTCTCCCGTGTGCGACAGACTAAATGATTCATAATGAATATGAATACCAGAGTTCACAGATGAGCAATAATCTCAGTTCCTGTTAAAGTG
ATGTCACTGTCGGCTTGTAGGGAGAGTCTGTGAAGGAAAATATCAATGTTAAAAGAAGATAAATTAAAGCAATAAGATGAACACCAACCTGAAGATTTTAGATCGCCTTTCT
CATCGAACTTGTATACGCGAGCATTAATAATACTCAGATCTGTCTGACTTGATTCCCTGTAAAAAAAAAAAAATGGGCAATTAAAATATTATAGCTTCTATTCGAATATTTCA
ACTTACACATTAATGGCCCACCAGGATTTGACGTGCAAATTCTTGCCAAGGAGCAGAATAAAACCAGCTATGGCCAGGCCCAGAATGCAAATTCCACTCCAGAAGTCAATGA
TGAAGAAATGCTCTCTCACTCTGGCCATTATATACACGACTATTTCTCGGATCCAGACCGCAAACACGCCCACTATGTAAGCGCTAAAGGCATAGAAACTGCAGGAATATAA
AAGATAAATAGGTCTATAATAAAAACAGCTTTTAGGAAATCTTACTGCGTCTTGACGATTTTCAGTTGAGTGCCCAATATGTACTGAGTATCTGGTGGCAGTTCGTAGATGG

AAAAGTACGCCTTGAGACAGCATACAATGCTGAATATCAGGAGGAGTGCAATGGGTCCGTAGTAGTAGATCAAAATACCCCAGGTGCGAACTGAAAATGCATAGATAATTCG
AAATTAGACTTTGTTGTACTTTCTTATTCAACCTACCATCGAACCAGCAGTAATCTTCTCCCATGCCTGGCTTGAAGTGCTTGGGCAGCTTGGAGTCCTGGGCTAAGGAGAC
CAAAAACCGAAGACCAATCACCGCCAGAACTGCCAGAGCAAGGGACAATCGACGGACGGAGCTCCTGACAGCCTTTTGTTTAAATTTGAGCAGAAAGTCAAGACTGCAAATG
GCCAGAAAGACGAAGGAGAGCATGATGCAGAAGTAGGCCAGAAATCCTTTTGGGATGCATAGAACAAAATAAGTTATCCAGTTGAAGTAGACAAAATTACTCAACCTATGTT
TCTGCAAGCCACATGCGAGAGCTTCATGGGATTCTTCAAGGCCAAGTAGACCAACAGGGAGTAACCAACGATCATGGACAGCAAGTAGTAGATGATTAACTGGCCATAGTGG
CTTTTACGGGCATCTCTCACAGATCCAAGTAGGGAGAGTATGAAAATGTTTATAATGATGGCTATAATACTGCCTAGAAGGGAACAAGAGTTAGAATCTCTTCGAAAGTATT
CAAACAACACGACTCACAAATGGCGTAAATCCAAACCCGGTACCCCGTTTGAAAGCGTTCGCAGTTGAGCGGAGTGAGCTCCCACTGCTCCGGATTGTGCTCCAACGGGCTA
AAGCAATACTCATCCGTTGACCACAGACGATTGTCCCGCCGCAGAGTTCCATTCTCGAAGAGGTCCCATTGCCAGAAGTTATCGTGCTTTTTGTCCACGAATTTGTTCCGGC
AACCCAGCTCTGTGCGCACCACAAAACGATCGCGTATGCGTACTTGATCCACAGTCCTGTTAGCGTAGGTGAGCTCCACGTGAGTGGACTCGCGGACTTGGTGTGGCATTGT
GCAGTTCCAGTGCTTGGCATCAAAGACGAGGTTCTCCGGACAACAGAAACTGATGCAGGGCTTCAGCAGACAGACGCAGCCGCGCAGATGTTTCTTGGCCCGGTACTCCACT
CCATCGATCACCTTGAAGGAGTACTCCGCCATCAGATGCGGTGGGACCACTACTCCAGCGTATGAGTAGGAGCCATCCTTCATCCGCAGTCCATCCGTAATGTTGACCGTGT
GGGCGTAGGCACAGGGATGGTTTATACTGGCGATCTCCACTGACTTTGCGCCCACGGAGAGCCAAATAATTAAAAGAATTATCAATGGCGAGACCATGGCTAACTGATCCAC
GCTTGTTGGATTTTTGGTCGATTCGTTCGTCGGCCTGACGTCACACGCTGTTCAGCTCGATAATTAAGTAAACAAATATGAGAAGCGATCGCCGTTCGCTTTGAGCAATGCA
CGCACACTCGCCGAGTCAGAAATCTGCACCAACAACAATGGAAACCCAAAATTGACAACGAATGTGACGCGGGTCTAAAAACGGCTGCGATTGCGACAACGACACCGACAAC
ACCAAAAAGAACAACCATAAACATAAAAGTTTGGTGGAGAGTCTCTCCTAGCTAGTCGTCTTGCCTCTCTCTAAACGACAAACCGGCAACCGCAACGGAGCTCCATGCTGGC
CAAGAATAAAAAACAGACTAACTGATTAGCCGTTCCCAAAACCAGACGCGACCACAAGCAGCTGGCAAAAAGAGCGATTGGGAAGACTCACAACCCGTCCCGCCCCTCCTCC
TAGAAACCGGCGTGGAATCTCAACCAGCTCGGAACTTGCACATGCATACCTTAAGTTGTAAGCATATAATCCATTCGAGCCATTTCCAGACTTATCTTGAGATAGATATAAA
AGTTAAGAATATTAGATTAGATTATTTCCATCCATTTTCTATAATTTATTGCCTTTGCGACCGTAAAGAAAATCCAATTTTTAACCGACAATCCAAATGGCCCAAACGAATT
TACCACAGCTTTAGTTTTTAAATATTTAGGTGCATTTCCAGGCTATACAAAGTTTGTCTACATGTAGGTGTGTGACCTGGCTTGATATCGTTTACGGTCTTCAAGAATATTC
GAAAAGATTCGTTGATTTTTGACAAACAGATCCGCTGAGGAAAAAACGACGTTCATTTTATCTTTGGCTTTATTAAGCTTGAAATACAATTATAATTCTTGTATTGTGTTTT
ACATTAGTTCAGAGATAATGCAGATATGGAGAATGGTACATTTATGCAAAGAGGGTATACAATCAACACCAATCTCGTTTTTAACAAGGCACATCTGTATGTTACGTATTTT
TGCTGGATGTTTACCTTGGACTTAGGCTAATGGAGAATTACAAATTTGATCTCTCAGCTGGGTTCGGAGAACCTTACAATTTTGCTTGCATAAACAGTAGATATGGCTCTGC
ATAAATTTCTTATACAACTAATCAGGTTGTTCGTGGTCACATGGCATTAAAAAATGAAAGAGAATTGGGCGGGGTGGGGTGGGTAAGGGTTTTCGTGGTAGTTCTGGGGGAT
AGGGACTTTCTCTTATAGTAGTGTCTTTCATAAAGTTCTATCCGTATGTGGATGCTACTTCGCGAATTTA
(SEQ ID NO: 163)

Exon: 3214..2258
Exon: 2201..2010
Exon: 1950..1717
Exon: 1658..1502
Exon: 1442..1239
Exon: 1176..1097
Exon: 1035..1001
Start ATG: 2897 (Reverse strand: CAT)

Transcript No. : CT33414
TCGTTTAGAGAGAGGCAAGACGACTAGCTAGGAGAGACTCTCCACCAAACTTTTATGTTTATGGTTGTTCTTTTTGGTGTTGTCGGTGTCGTTGTCGCAATCGCAGCCGTTT
TTAGACCCGCGTCACATTCGTTGTCAATTTTGGGTTTCCATTGTTGTTGGTGCAGATTTCTGACTCGGCGAGTGTGCGTGCATTGCTCAAAGCGAACGGCGATCGCTTCTCA
TATTTGTTTACTTAATTATCGAGCTGAACAGCGTGTGACGTCAGGCCGACGAACGAATCGACCAAAAATCCAACAAGCGTGGATCAGTTAGCCATGGTCTCGCCATTGATAA
TTCTTTTAATTATTTGGCTCTCCGTGGGCGCAAAGTCAGTGGAGATCGCCAGTATAAACCATCCCTGTGCCTACGCCCACACGGTCAACATTACGGATGGACTGCGGATGAA
GGATGGCTCCTACTCATACGCTGGAGTAGTGGTCCCACCGCATCTGATGGCGGAGTACTCCTTCAAGGTGATCGATGGAGTGGAGTACCGGGCCAAGAAACATCTGCGCGGC
TGCGTCTGTCTGCTGAAGCCCTGCATCAGTTTCTGTTGTCCGGAGAACCTCGTCTTTGATGCCAAGCACTGGAACTGCACAATGCCACACCAAGTCCGCGAGTCCACTCACG
TGGAGCTCACCTACGCTAACAGGACTGTGGATCAAGTACGCATACGCGATCGTTTTGTGGTGCGCACAGAGCTGGGTTGCCGGAACAAATTCGTGGACAAAAAGCACGATAA
CTTCTGGCAATGGGACCTCTTCGAGAATGGAACTCTGCGGCGGGACAATCGTCTGTGGTCAACGGATGAGTATTGCTTTAGCCCGTTGGAGCACAATCCGGAGCAGTGGGAG
CTCACTCCGCTCAACTGCGAACGCTTTCAAACGGGGTACCGGGTTTGGATTTACGCCATTTGCAGTATTATAGCCATCATTATAAACATTTTCATACTCTCCCTACTTGGAT
CTGTGAGAGATGCCCGTAAAAGCCACTATGGCCAGTTAATCATCTACTACTTGCTGTCCATGATCGTTGGTTACTCCCTGTTGGTCTACTTGGCCTTGAAGAATCCCATGAA
GCTCTCGCATGTGGCTTGCAGAAACATAGGATTTCTGGCCTACTTCTGCATCATGCTCTCCTTCGTCTTTCTGGCCATTTGCAGTCTTGACTTTCTGCTCAAATTTAAACAA
AAGGCTGTCAGGAGCTCCGTCCGTCGATTGTCCCTTGCTCTGGCAGTTCTGGCGGTGATTGGTCTTCGGTTTTTGGTCTCCTTAGCCCAGGACTCCAAGCTGCCCAAGCACT
TCAAGCCAGGCATGGGAGAAGATTACTGCTGGTTCGATGTTCGCACCTGGGGTATTTTGATCTACTACTACGGACCCATTGCACTCCTCCTGATATTCAGCATTGTATGCTG
TCTCAAGGCGTACTTTTCCATCTACGAACTGCCACCAGATACTCAGTACATATTGGGCACTCAACTGAAAATCGTCAAGACGCATTTCTATGCCTTTAGCGCTTACATAGTG
GGCGTGTTTGCGGTCTGGATCCGAGAAATAGTCGTGTATATAATGGCCAGAGTGAGAGAGCATTTCTTCATCATTGACTTCTGGAGTGGAATTTGCATTCTGGGCCTGGCCA
TAGCTGGTTTTATTCTGCTCCTTGGCAAGAATTTGCACGTCAAATCCTGGTGGGCCATTAATGTGGAATCAAGTCAGACAGATCTGAGTATTATTAATGCTCGCGTATACAA
GTTCGATGAGAAAGGCGATCTAAAATCTTCAGACTCTCCCTACAAGCCGACAGTGACATCACTTTAA
(SEQ ID NO: 164)

Start ATG: 318 (Reverse strand: CAT)

MVSPLIILLIIWLSVGAKSVEIASINHPCAYAHTVNITDGLRMKDGSYSYAGVVVPPHLMAEYSFKVIDGVEYRAKKHLRGCVCLLKPCISFCCPENLVFDAKHWNCTMPHQ
VRESTHVELTYANRTVDQVRIRDRFVVRTELGCRNKFVDKKHDNFWQWDLFENGTLRRDNRLWSTDEYCFSPLEHNPEQWELTPLNCERFQTGYRVWIYAICSIIAIIINIF
ILSLLGSVRDARKSHYGQLIIYYLLSMIVGYSLLVYLALKNPMKLSHVACRNIGFLAYFCIMLSFVFLAICSLDFLLKFKQKAVRSSVRRLSLALAVLAVIGLRFLVSLAQD
SKLPKHFKPGMGEDYCWFDVRTWGILIYYYGPIALLLIFSIVCCLKAYFSIYELPPDTQYILGTQLKIVKTHFYAFSAYIVGVFAVWIREIVVYIMARVREHFFIIDFWSGI
CILGLAIAGFILLLGKNLHVKSWWAINVESSQTDLSIINARVYKFDEKGDLKSSDSPYKPTVTSL*
(SEQ ID NO: 165)

Name: mth-like 9
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384809
CGAGCCATTTCCAGACTTATCTTGAGATAGATATAAAAGTTAAGAATATTAGATTAGATTATTTCCATCCATTTTCTATAATTTATTGCCTTTGCGACCGTAAAGAAAATCC
AATTTTTAACCGACAATCCAAATGGCCCAAACGAATTTACCACAGCTTTAGTTTTTAAATATTTAGGTGCATTTCCAGGCTATACAAAGTTTGTCTACATGTAGGTGTGTGA
CCTGGCTTGATATCGTTTACGGTCTTCAAGAATATTCGAAAAGATTCGTTGATTTTTGACAAACAGATCCGCTGAGGAAAAAACGACGTTCATTTTATCTTTGGCTTTATTA
AGCTTGAAATACAATTATAATTCTTGTATTGTGTTTTACATTAGTTCAGAGATAATGCAGATATGGAGAATGGTACATTTATGCAAAGAGGGTATACAATCAACACCAATCT
CGTTTTTAACAAGGCACATCTGTATGTTACGTATTTTTGCTGGATGTTTACCTTGGACTTAGGCTAATGGAGAATTACAAATTTGATCTCTCAGCTGGGTTCGGAGAACCTT
ACAATTTTGCTTGCATAAACAGTAGATATGGCTCTGCATAAATTTCTTATACAACTAATCAGGTTGTTCGTGGTCACATGGCATTAAAAAATGAAAGAGAATTGGGCGGGGT
GGGGTGGGTAAGGGTTTTCGTGGTAGTTCTGGGGGATAGGGACTTTCTCTTATAGTAGTGTCTTTCATAAAGTTCTATCCGTATGTGGATGCTACTTCGCGAATTTATAGAT

ATCGACTTTAAATGTTGATTACAGAATAATCTCATGATAGGAATATGTAGTCCTAAAATATGAATAAGTGATCTTTATCTACAGAAAGGAACGCAGTGTTCGCAATAATGCG
AGGATCCGATCGAAATGAATATAAATGCGATTAGTCCTTCCTGATCTGCACGCATCCCGAGGTATTCAATGCTCTGCCTATGGCGGCAACTGCGAGTGCTGTGATCAGCGGA
AGATGGTGGTTTTCTGAGGATCCACGCTCGAGCTGATCACTAGCGGCTTCTCCACCGATGGCTTCTCGTGCAGCGACAGATTGGCCACGCTGCTCGACGTGGTCTTGGTGGA
GTACTGGCTCTGACGTTGGTTGCTGCCATCGCGAACACTCGAGCATCTGTGTTTGTGTGAAGTGGTGGAGTGGTTATGACGTCCATTGCGTGGGCAGCGTGGAATAGACAAT
GGCGGTAAATCTGATTAGAATTGGTTCGTTTGAGTGCTTGGATTTTAAGTTAGGGAAAAGCAGACACATTAAGCGATCCTACAGGATATAAACAGGTTACCTTCACCAAGAC
ACAAAAGTGGATGGTCATTATGATTAGTTTTCGAGGGGTTAGAGGTTATCAGGGTTATCTAATGGTTTCTTCCAACTGGGGTTAACTACAATGTCAACTACTTCTCCATTTC
CAGAGGATTACCTGTGACCATTTTGATTAGATGGCTATCTTCGGGGTCGGACTCGGGTCGGTGGCATTGCTACCGGCCTTGTCAACAGCCTTAAGGTGCAGCTTTCTTGGCA
GCACATTCGCGTCAGAAGTGATGGAAGTCGAGGTAGTGGTCGAGCTCAGGCCATTCGAGCGCCTTTTCTCCAGGAGTCTTCAATGAATAGCGATCGATGATGATCAAATTGA
TGTTGAGACACATGCACATATGGACACACATGGTGGATGCAAACATAAGTTTGTGTTTAGATCAGGGCAATTAGGAATGGTACTTCTGTTTAGTAGCTACAAAGACGCGACC
GGACTGAGTCACCGAGCTAGGCGGCTGGAGATTCAGTGCTGTGTTGGCTGTCAGTGTTCGGTTTAGAGTTAGTCTGCAGGCAGTGTCTGTTGGTGTTGAGCTTGCTTTTAAG
CAATATCGAAAGCAGTCTGACAACGAAAATAAGTACAACATATAAGGTTTAAAATAATCACAGATTGCGCGCTGTCAGAACTCGAAGTGCATATGTATTCCCATAATCTTCA
ATGGGTCTTCCATCTTAGTGCAAGAGATTTAAGGAACGTGGAATTAGACAGGAATCATATAACCCGAAATTTTCAATATTCTTGATTTAAGCGCGGGTTAGCAAAGGCCACA
ACAAACGAGCAACCGAGCTAAAGGAGACCTTCTTACTGAAATCGACGAAATCGATTTAGACTTCGTGTGAATGAATACCACTGGTGGACCCATTGGCAAATGCGTTGCGGAT
CACCACCAGCTTGCGGATCTGGTTGGAGGCGTCCATAAACCGCTGCTTCATCGAGAACGAAGTCGGTCCCATCGTTACGCTCTCAGTCTGCCGCAGACTGCGCACACGTATC
GTTCTGAGAATAGTCGGGGTTACAGATATGATGGTGGACAGAGTCAGAGACACGACACACAAGAGGGAGTCACGAGTCGTGTGAATGGACAAACGCCCAGTGGGCGGGACAG
GACATAACGTGGAGTTAAGAATAGAGCAACACTTACCTGTTTGTGATCAAGTGCTTGACCTTCTTCTTCATTACAAACAACATGAAAATGAGAAAGCCCTGCATCGCATTGG
CCAGATCCGAGATATAAAAGAGCTTAGACCATCCCTTGTCGCTGCCCACGAAGTACGAGATAAGCTCCGTGAGCCAGGTGATGCCCATTATGAGAAAGAGGCGTAAGAACAA
GCCGAATCTATAGTTGGACCAAGCACGATAAAAACGTCTTTAGGGGAAGGTGGATAAGAAATACTGAAAGGATGACACTCACTTGTCCTTCTCGGTTCGTAGGTTCTTGGTG
CTGTTCTCACTCGCAATTATCCGGGCCATCTCGCGCTGCACGCCGTGAATCTTAATCGCCGTCATTATGAACATGATGGTGTTGGCAACAACGATCGCCAGAATGGGACCGT
AGAAATAGATCATGGCTGCCCAGTTGCTCATATCCAGCCAGCAGTATACACCATCACCAATACCTGGCTTTAGGTTCGCGGGCAAATTGGTAAGTTGCTGGGCGCAATAGGT
GAAGGCTAGGAAAACTAGGGCAATACCCCAAGAGTACAGCGAATAAAACAGGAAGCGTTTCTTCTCCTGGAATCGGTTAATTCCTCTCGTCCCTCGGAAATTGTGCCACAAG
TCAAAGCTAATGACACTTAGCCACATGTAGGCACCCATAAAAAAGAAATATGCTGTGTATCCAAACACTTTGCATGTAACACCAGTGGCATCAACTTGATAGAGTTGCACGT
AGCACAGCGAGGAGTAACCCACGGAGAGCCCTATTAAATAGCAGACCAAGCTCTTTCCGTGCTGGTTGCGGAGCTCGGGAATCAATAGGTAAACGGCTATGGTAAGCATCAT
GAAGGGAATGGAGAACATCATGGCTAGTGCAAGGAATTTCGGTCGGATTAATAAAATGCTGTTTACAAGAAGTAAAGGTGATCGTTTCACTTACCATACGAATTGATAATCT
TTACCGTGCTATGATCGTCCTGCATGTCACAATTGGCCGGATGTATGGTGTAGAACATGTCCGATTCGTTGACGTATGTCGGAACCATGCAGAACTCGTTCTTCCTGATTAG
CAGTTTATCGTCCACACGGAGCATGGATCCGTTCTCAAAGAGTAGATAATCATCCCAATTGTTGGTCTCTGGCTGCAGAGAAAACATCTTGGGGCAGGGTCTGAAGCTTTGG
ATAGCAAACTGCGTAAACATGTTGACCAATATCGTAGACTCGTTGCGGAAGGTTATATTGAGCATAGGCGGATCTGGCCACGTTCGGTTGTCCGAGTGGTCTATTATGCATC
CGTCAACTTCAGAGTTAAAGACCTGCCGCCACGGGCAGCAGATATTCAGGCAGGACTTGAACTTGCAGACGCATCCCCTAACGTGTTCCCGAACCTCTATGCGCTCCGTTAG
CGAGCTGTGTATGTAGTCGTACGTCCCGACCAATTCTGCGGGAACTATTGTGCCGTAGTAATCATAGCTTCCGTTGGGAAAGAGCCTATGGCCCGTCAGATTCACTGTATCG
TAGAAATTACAGGGCACATTTGGATCGTCTATGGTGTGGTAGTGGTCTCTCTCGGCCATAACGTAAGTGCCGAAGGGAATCCCCGGAATCAACCGGAACACGACCAAGCACA
GGACGCCCAGCAATGTAACTCCGCAGTATAGGGATCCTCCAGGCTGATGAATTTTCTTCGGCATAATGAAGGTTCTGAAATTCGTAAATCTTTATTAAATATTATTACAAAT
ATTTGTTATTATTCTTTAACGTTTATTCTTTTAACGTTTAAATATTCAGTGGAATATAGTTATTACAGTTAGTCATCAGTCGCAACACGCAGCAATTAAGCCTTAGCAAGTG
AAGTCACTAATTAAGACCTATTTGCCGATTAGTTAGTCGTTTAATGTTTATCTAAGTGAGTTCACTAATTTGGGCAATAGAGCAACATCTAGGGACCCATTAGCGGGGACGA
GTGGTGAATGATCTATGCGGACAACAGGTTGTCGAAAAGAATTAAGGAGGCGAAACAGGATAGCGTCACAAGCAATTAGCAGCCAACGTGTCCCATGGGGAGGAACCAGGGG
CGGACTACCCACGACCTGAGCGCCATTCTCCGCGGGACATCAATCATAAGGTCAACCACGCAGGGAGTGCAGACTGAGCAAACATTTATGGATTCGTGTTAAATTGATAGTA
ACTGCCATTGTCTGTCAATAGAGGTGACAGTCTCCCCGAACTCCACTCCAAACCAATCCCGATCTGCCCACTTTTCAAGCGCTGCTAACGAGCCCGTCCATAAATTCATTAC
CAGTGCGAAAAGTTATACCGCTCGCTATCTTATCTTGTATCATTAGCCTTTGGGGACCAACTAATTCATTGATCCGCTGATTGTTTGCCTTTGGGGGAGAAGCAGAAGGGTA
ACCTGCGGCATTGTCAGGAGTGTGGGCCACTAAATTAGTCCGAATTTATGGGATGCGAACGGGTGGAGAAGGACAAACACATTAAACTCCGCACAGTTGGCTAATGGGGAAA
ATGAATACTGGCAGTGGGTGGGGAAGGGAGAGCCAGAGCACTTATCTTGCGGTGGGAAATTAGCATAATTAATCGCTGAGCACATGCACTTGGGAAATTTCAGCGGACCAAA
TTCAATCGTATATATCTGCGCACTTTAAATGTGCGAATGAGTGCTGCGGATGTTTG
(SEQ ID NO: 166)

Exon: 4208..3455
Exon: 3344..2771
Exon: 2695..2501
Exon: 2355..2207
Exon: 1242..1001
Start ATG: 4208 (Reverse strand: CAT)

Transcript No. : CT33415
ATGCCGAAGAAAATTCATCAGCCTGGAGGATCCCTATACTGCGGAGTTACATTGCTGGGCGTCCTGTGCTTGGTCGTGTTCCGGTTGATTCCGGGGATTCCCTTCGGCACTT
ACGTTATGGCCGAGAGAGACCACTACCACACCATAGACGATCCAAATGTGCCCTGTAATTTCTACGATACAGTGAATCTGACGGGCCATAGGCTCTTTCCCAACGGAAGCTA
TGATTACTACGGCACAATAGTTCCCGCAGAATTGGTCGGGACGTACGACTACATACACAGCTCGCTAACGGAGCGCATAGAGGTTCGGGAACACGTTAGGGGATGCGTCTGC
AAGTTCAAGTCCTGCCTGAATATCTGCTGCCCGTGGCGGCAGGTCTTTAACTCTGAAGTTGACGGATGCATAATAGACCACTCGGACAACCGAACGTGGCCAGATCCGCCTA
TGCTCAATATAACCTTCCGCAACGAGTCTACGATATTGGTCAACATGTTTACGCAGTTTGCTATCCAAAGCTTCAGACCCTGCCCCAAGATGTTTTCTCTGCAGCCAGAGAC
CAACAATTGGGATGATTATCTACTCTTTGAGAACGGATCCATGCTCCGTGTGGACGATAAACTGCTAATCAGGAAGAACGAGTTCTGCATGGTTCCGACATACGTCAACGAA
TCGGACATGTTCTACACCATACATCCGGCCAATTGTGACATGCAGGACGATCATAGCACGGTAAAGATTATCAATTCGTATGCCGTTTACCTATTGATTCCCGAGCTCCGCA
ACCAGCACGGAAAGAGCTTGGTCTGCTATTTAATAGGGCTCTCCGTGGGTTACTCCTCGCTGTGCTACGTGCAACTCTATCAAGTTGATGCCACTGGTGTTACATGCAAAGT
GTTTGGATACACAGCATATTTCTTTTTTATGGGTGCCTACATGTGGCTAAGTGTCATTAGCTTTGACTTGTGGCACAATTTCCGAGGGACGAGAGGAATTAACCGATTCCAG
GAGAAGAAACGCTTCCTGTTTTATTCGCTGTACTCTTGGGGTATTGCCCTAGTTTTCCTAGCCTTCACCTATTGCGCCCAGCAACTTACCAATTTGCCCGCGAACCTAAAGC
CAGGTATTGGTGATGGTGTATACTGCTGGCTGGATATGAGCAACTGGGCAGCCATGATCTATTTCTACGGTCCCATTCTGGCGATCGTTGTTGCCAACACCATCATGTTCAT
AATGACGGCGATTAAGATTCACGGCGTGCAGCGCGAGATGGCCCGGATAATTGCGAGTGAGAACAGCACCAAGAACCTACGAACCGAGAAGGACAAATTCGGCTTGTTCTTA
CGCCTCTTTCTCATAATGGGCATCACCTGGCTCACGGAGCTTATCTCGTACTTCGTGGGCAGCGACAAGGGATGGTCTAAGCTCTTTTATATCTCGGATCTGGCCAATGCGA
TGCAGGGCTTTCTCATTTTCATGTTGTTTGTAATGAAGAAGAAGGTCAAGCACTTGATCACAAACAGAACGATACGTGTGCGCAGTCTGCGGCAGACTGAGAGCGTAACGAT
GGGACCGACTTCGTTCTCGATGAAGCAGCGGTTTATGGACGCCTCCAACCAGATCCGCAAGCTGGTGGTGATCCGCAACGCATTTGCCAATGGGTCCACCAGTGATTTACCG
CCATTGTCTATTCCACGCTGCCCACGCAATGGACGTCATAACCACTCCACCACTTCACACAAACACAGATGCTCGAGTGTTCGCGATGGCAGCAACCAACGTCAGAGCCAGT
ACTCCACCAAGACCACGTCGAGCAGCGTGGCCAATCTGTCGCTGCACGAGAAGCCATCGGTGGAGAAGCCGCTAGTGATCAGCTCGAGCGTGGATCCTCAGAAAACCACCAT
CTTCCGCTGA
(SEQ ID NO: 167)

Start ATG: 1 (Reverse strand: CAT)

MPKKIHQPGGSLYCGVTLLGVLCLVVFRLIPGIPFGTYVMAERDHYHTIDDPNVPCNFYDTVNLTGHRLFPNGSYDYYGTIVPAELVGTYDYIHSSLTERIEVREHVRGCVC
KFKSCLNICCPWRQVFNSEVDGCIIDHSDNRTWPDPPMLNITFRNESTILVNMFTQFAIQSFRPCPKMFSLQPETNNWDDYLLFENGSMLRVDDKLLIRKNEFCMVPTYVNE
SDMFYTIHPANCDMQDDHSTVKIINSYAVYLLIPELRNQHGKSLVCYLIGLSVGYSSLCYVQLYQVDATGVTCKVFGYTAYFFFMGAYMWLSVISFDLWHNFRGTRGINRFQ
EKKRFLFYSLYSWGIALVFLAFTYCAQQLTNLPANLKPGIGDGVYCWLDMSNWAAMIYFYGPILAIVVANTIMFIMTAIKIHGVQREMARIIASENSTKNLRTEKDKFGLFL

RLFLIMGITWLTELISYFVGSDKGWSKLFYISDLANAMQGFLIFMLFVMKKKVKHLITNRTIRVRSLRQTESVTMGPTSFSMKQRFMDASNQIRKLVVIRNAFANGSTSDLP
PLSIPRCPRNGRHNHSTTSHKHRCSSVRDGSNQRQSQYSTKTTSSSVANLSLHEKPSVEKPLVISSSVDPQKTTIFR*
(SEQ ID NO: 168)

Name: mth-like 10
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384577
CCAGACATGCTTCCTCGATGGGGCTTAACGTCAGCGGAGCAGCGATGTCAGTGGCAGCCATCATCTGAACCGAATGACAGTGGAGGTCGCTAGAAAAGGGGTTCCATTTCCA
CTAGTTTATACTACCGAACTGATACTTAAATCAAGGCAGTTAAAAATTTAATATGAAACCGGAAATGGCTTCAACATCTCGTTAACGCTCTGCCCAATTCTACAGTCATCCG
ATTTTAATGAAGCAAAATTTACTGCCATGCCATGCAAGAGATATGAATCACCGCATGGGAATAGCCCACAAAGGACCCTTACCCCATGCACTCTATCATGAACACCGCACAC
ACTTTTAATCTGCGACTAATTGCTGCTCCGTATTGAATTGATTTCGGTTGCCCCAAAGTGCCACATTCCAAATCTCCGTTGATTGGGGAAATCCCCGAAAAACATTCGAACC
ATTCGTCCCTGGCGAGCACAGAAACAGCGTCTCTGATTCTGGCAGGAATTGGAGGGCATAGGGCGGACCTTGTGGCGGTTTGTCGAGTGCACAAAGTGCAAGGTTGCGTGGC
CATTCATTTCGCAGCCGATGGGTGGAAGTGGGCAGCGGTGGGTGGCGCCCGGAAGGAAGAGCATCAGTAATCTCTAAGTGCGCCTTGTTCGCATTCGTACTCGTGCCAGCAT
TGCATAAGTGTCAGTGTCAACGGGGCGGGGATGTGGTTGTGCTTGGTGGTTATGTGGGTGTTTGTGTGGTAGTAGTGTCCTTGTTGGCAGACAAAAAGACTGATGATGATGG
CATTTTGTTACTTTGCTACCTTCCTTTGTCGCCCGTCGCACGGCAGCTCCTTTAGTATGAAATGAGGCGACCAGGCAAGCGCCTCGTCATCGAAGATCACCCAAACAGAGTA
CGGACCAGGACACGGCAACTCGCACAGGCACAGGAACAGGTGCAGGCACAACAACTGGCAGGAGCATCGGTGTCGGCAAGGGATACTGGCGCGGGATCGAAAGCATGAACAG
GGACAATCTGCAGCAGTGGTGGGAGAACTCGTATCGGCGCCAACATCCAGAGCCCACCGATGACCTCGGCCTGGACTCCGCGGAGCTGCATCTCGCCCTCCAGGAGCCCAAC
CAGCTGCCCGCGGACTACGACTATGGCAACTTTAGTCTAGGCAATCCGTATGACGTGGACTCAGAGCACTCCATTTCCCCGCTGACGCTGCTCCTGCTCGCTGTTAGCTATG
GCCTGGTGGTTTTCGGCGGCGTCGTGGGCAACTCCACGCTCGTTCTGACCCTCTGCTCCGCCTCTTCGGTGCGTTTGCGTAACCCGCTGCTGCTGGCCGTCTGCATTGCCGA
TCTCCTGGTCACCGGCATCTCCGCTCCGGTCACGCTCCTCAATCTCGCGATGAACCGGAGGACGCGATCGCTGCCCCTTGTGCTGTGCAAGGTCATACACTACGTACAGGTG
AGTGGGTCAATCAAACAAGCTTAAGTTAGCCAAGTAACTAAGTTAACCCCACGCTATTCAGGTCATGCCCGTTTCGGCCAGCACTATTTCGTTTTTCATGCTCTCCCTTGAC
CGCTACGCCACAGTAAAGCATCCGCGATTGGCGCAGCTTCGCCAGCGGCGCTACTTGCACGTCTCCCTGGCGCTGCTATCCTGGCTCGCCTCTGCCGCCATCAGCACTCCCT
TTCTGTTTGCCTACAAAATCATCGCCAAGTCGATGGTGGTCAAGGGCGGTGGGGCGGCGAACACCACGCCAAACCCGGTCAGCATCAGCTGCACCTCCGACCTGGGCGCCAA
CGCCATGTTCATGTCCTTCATCATCTTCCACACCATCGCCGTGTTCGTGTTGCCTGGCATTGGTGTGCTACTCAACCACTACGGCGTCCGACGCAAGCTCTGCGCCCTGTCG
CTAACGGCTCGAGCTGCCCATGGAGAGCTGCCGCTGCCCATTCCCATCCTGCGACGCCAGACCCACATGGTCATTGTGACGGGCTGCCCCAACGCCCAGCAAGCGGCCTGTG
GTGGGGCACCACGGCGGACGACACCTCGAACGGGAACGGCACGGGCACCGGTGGCGGTCCCATGGCCGTCAGTCCGGGCGACATACAGCTTCACACACTGCAGCCGCGCCA
GCCAGGCTCTGCCGGATCGGCCCTCGAACCGGGCTCCTACCGCTCTAGCAATCCCATATCACCCAGGTAAGTCATAAATATTGTGAAGTTATCAGCAAAATCGTTGTAAACA
GGATGTATGCTTGATGTAGTCAAAAGCTATTAGTTTATGTTAATCATACGCCCCGTTGACTTTGGTCAGAATATATTTCATCGAATGGCAGTACGTCATTTGGTATAAGTTA
ACTAACTTGCAAACCCACTTAGGGCCATGAGGGAGATTCGCGCCCACTCGCAGCGCCAGCGCATCAACCGAGCAGGCAGGGGACCAGCGACACCCGGCATCCCACTGCCCCA
AACCTCCACGCTGCGGTCGCGAAGGCACCTGGCTAACATGCTCATCGCCTCGGCCGTGATCTTCATTGCCTGCTGGGCACCGCACGTCTTCTGCATCTTCTACAAGAACTTC
GGCAACAATCAGCAGTGCAGCCAGACCTCTGTTTACTTCAGCCTTCTGCTAGGTAAGTCTGAATTGGCAATCATTCCCAATCCGAATCCGAGTCCGAATTGGAACTGACTAG
TTGATCGCGTTGCAGGCTACTTCTACTCGGCCATCAGTCCGGTGATCTACTGGGCGCTCAACCACAATTCGCTTCGACAATCTCCCTGCGCGCCCATCATCCGCCTGCGCTC
CATGCAGAACTTCCTGAGGTCGCGCTTCCGGACGCACACCGCCCCGCCGCCTCCCTCGTCCACCAACGAGGCGGCACTGGGGGCATTCAACCCCAAGCTGATCAAGCTAACA
CCGAAACAGTATAGAGCTCAGGCCTCTTCGCATTATCTCTACTAGTGTGATACTAAGCTAAACTTGATATACATATATACACATTTATATACACGCATATATAGTATATTTT
TTACCTACGACCCCACCCCCCCCCTGACCCGTCGATGCCCTCTTAGTACCCTCTAGTATCGAACTCCGAACCCGAACTTTGACCCAACCGAACGAAACTATCTAGATGTTTT
TCAAAAGCTCTACGCAACAATTGCATAATACGAAAGATACATGCTACATATACACATATATAGAGTACATACTACCTTGCATAGCAGTGTTAGTCGACGCTTATTTACACGA
TGTTGTGCGCCTGAAGCTATGCAATAAGGACACCGGAGTCCTGCCGCAACTCTTGCCTATAACGAGGATAGCACTGGTAACGAAAGCAAGAAAAAAAAAAAGGAAAAAAAATA
CAAAAATAACGAAACCCTTCTAGTGACTTGTTCGAATTGTTGTTTAGTCTTTTAAGCTGAAATATTGTACTATGTTCCTTCTCATAGATAAGCAATAAGTGAACGCTTTAAG
CCAAGTTAAACAAACGGAAAACGAATCAAACTAAACTAAAACTCGATCTAAAACTTTTAAACGATTTAAACGTAAACTCGAAAGCATACCTAACGAAACATTAATAATTATA
TACATACGCGACTACATACAAATAACAACTAACAAGTGAAAACTAAAAACAAAAAATGAATGTAAATGTGAGCAAAAAAACAAAAGGATAACATTTAGTATTGTATACATAT
TCTTCGACTTAATTCTTTATACACTTTTTCTAACCATTTTTAATATCGTTCAGTAAACTGAACACGGTGTATAGGGCAATTCAAGCTGCTGGATATATTCGCGTTTCAGATG
CCAGAGCCACCCAAAGGATCGGACTCTGGATCGGGTTTCATTTCACTCCATTCAGTGGCGACCGTTTCCTAATTTGCCAAACTTTTGTACATGCTCCGCTGCGAGCGACTCT
CTCCTTACTCCACCTGGTTTTTCGTTCAATCCTTGGA
(SEQ ID NO: 169)

Exon: 1001..1453
Exon: 1518..2194
Exon: 2375..2628
Exon: 2704..2957
Start ATG: 1001

Transcript No. : CT33551
ATGAACAGGGACAATCTGCAGCAGTGGTGGGAGAACTCGTATCGGCGCCAACATCCAGAGCCCACCGATGACCTCGGCCTGGACTCCGCGGAGCTGCATCTCGCCCTCCAGG
AGCCCAACCAGCTGCCCGCGGACTACGACTATGGCAACTTTAGTCTAGGCAATCCGTATGACGTGGACTCAGAGCACTCCATTTCCCCGCTGACGCTGCTCCTGCTCGCTGT
TAGCTATGGCCTGGTGGTTTTCGGCGGCGTCGTGGGCAACTCCACGCTCGTTCTGACCCTCTGCTCCGCCTCTTCGGTGCGTTTGCGTAACCCGCTGCTGCTGGCCGTCTGC
ATTGCCGATCTCCTGGTCACCGGCATCTCCGCTCCGGTCACGCTCCTCAATCTCGCGATGAACCGGAGGACGCGATCGCTGCCCCTTGTGCTGTGCAAGGTCATACACTACG
TACAGGTCATGCCCGTTTCGGCCAGCACTATTTCGTTTTTCATGCTCTCCCTTGACCGCTACGCCACAGTAAAGCATCCGCGATTGGCGCAGCTTCGCCAGCGGCGCTACTT
GCACGTCTCCCTGGCGCTGCTATCCTGGCTCGCCTCTGCCGCCATCAGCACTCCCTTTCTGTTTGCCTACAAAATCATCGCCAAGTCGATGGTGGTCAAGGGCGGTGGGGCG
GCGAACACCACGCCAAACCCGGTCAGCATCAGCTGCACCTCCGACCTGGGCGCCAACGCCATGTTCATGTCCTTCATCATCTTCCACACCATCGCCGTGTTCGTGTTGCCTG
GCATTGGTGTGCTACTCAACCACTACGGCGTCCGACGCAAGCTCTGCGCCCTGTCGCTAACGGCTCGAGCTGCCCATGGAGAGCTGCCGCTGCCCATTCCCATCCTGCGACG
CCAGACCCACATGGTCATTGTGACGGGCTGCCCCAACGCCCAGCAAGCGGCCTGTGGTGGGGGCACCACGGCGGACGACACCTCGAACGGGAACGGCACGGGCACCGGTGGC
GGTCCCATGGCCGTCAGTCCGGGCGACATACAGCTTCACACACTGCAGCCGCGCCAGCCAGGCTCTGCCGGATCGGCCCTCGAACCGGGCTCCTACCGCTCTAGCAATCCCA
TATCACCCAGGGCCATGAGGGAGATTCGCGCCCACTCGCAGCGCCAGCGCATCAACCGAGCAGGCAGGGGACCAGCGACACCCGGCATCCCACTGCCCCAAACCTCCACGCT
GCGGTCGCGAAGGCACCTGGCTAACATGCTCATCGCCTCGGCCGTGATCTTCATTGCCTGCTGGGCACCGCACGTCTTCTGCATCTTCTACAAGAACTTCGGCAACAATCAG
CAGTGCAGCCAGACCTCTGTTTACTTCAGCCTTCTGCTAGGCTACTTCTACTCGGCCATCAGTCCGGTGATCTACTGGGCGCTCAACCACAATTCGCTTCGACAATCTCCCT
GCGCGCCCATCATCCGCCTGCGCTCCATGCAGAACTTCCTGAGGTCGCGCTTCCGGACGCACACCGCCCCGCCGCCTCCCTCGTCCACCAACGAGGCGGCACTGGGGGCATT
CAACCCCAAGCTGATCAAGCTAACACCGAAACAGTATAGAGCTCAGGCCTCTTCGCATTATCTCTACTAG
(SEQ ID NO: 170)

Start ATG: 1

MNRDNLQQWWENSYRRQHPEPTDDLGLDSAELHLALQEPNQLPADYDYGNFSLGNPYDVDSEHSISPLTLLLLAVSYGLVVFGGVVGNSTLVLTLCSASSVRLRNPLLLAVC
IADLLVTGISAPVTLLNLAMNRRTRSLPLVLCKVIHYVQVMPVSASTISFFMLSLDRYATVKHPRLAQLRQRRYLHVSLALLSWLASAAISTPFLFAYKIIAKSMVVKGGGA
ANTTPNPVSISCTSDLGANAMFMSFIIFHTIAVFVLPGIGVLLNHYGVRRKLCALSLTARAAHGELPLPIPILRRQTHMVIVTGCPNAQQAACGGGTTADDTSNGNGTGTGG

GPMAVSPGDIQLHTLQPRQPGSAGSALEPGSYRSSNPISPRAMREIRAHSQRQRINRAGRGPATPGIPLPQTSTLRSRRHLANMLIASAVIFIACWAPHVFCIFYKNFGNNQ
QCSQTSVYFSLLLGYFYSAISPVIYWALNHNSLRQSPCAPIIRLRSMQNFLRSRFRTHTAPPPPSSTNEAALGAFNPKLIKLTPKQYRAQASSHYLY*
(SEQ ID NO: 171)

Name: GASTRIN/CHOLECYSTOKININ RECEPTOR like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384577
GAATGATAACTAATTCAATAAAACTACAGATTTCAATAACTAAAGCATACATGTATATGTATCTATGTATCTATGACTATGTATCTCGAATCCTCCGCTGAAACTCTAAACT
ATTCCCATGCAGTTTCTTTTCCTTCGTATCGAGTTTCAGAGCACGGCTTGTAGTGTTGCACGTTTTTCGTATTTTTTGGCAGCTCTGGCTGGTTTTTTTCATTTTCCTTTTT
TTTTTGTACTCCCATTTCCTGTCACCCTGCCCGGTGTAATTAAAAAAGTTTTCTCGACATTTGCTTGCCATGTGAGGCTCCAGGAACAAATGCGTTACATTGTGGACCATC
CGATTGTAAAATAAATTTTAAGTGCAATACGAAGTGCAAACTCAGCCGTAACAATAAGATTTTTGCAATAATATTTAACTTTTAATTGTAACTTTGGTAAAGTCAATGTTTT
CTTTTTCCAAGTTGGCCAATAAAAAGAGCATAAAGCAAAGCGAAAACTTTAGTGCGTACTTTTATTACAACATTAACTTCATGCTATATATTTTATTTAAATATTTTGGCTT
TTCTATCCATTGTTTTTCAGTTAATATCTTAAACAAACAATTTTTGATGCACTACAAATGCTGTTGATTTTAATAAACGATTTATAGCTTACATGGATTATTAGCTATGTAC
ATATAATTACGAATATTTTTAATAAAGGGCATACTTAATAAATTGAAATTCTGGAAAACACTCATTTATTTAACACTAGCAAATTAGCGCACAAATAATGTGCAAAGTTCGT
TAATCTTTAAAGTGTATTTGTGACCACAGCACAATTAAACGGTGGTCCGAAATTCTATATCTTCTAATTCCGCATTGTTTTTTTTTTTTTGTGAATTTTTCAGGTAACCTCC
TTGTTATTTTGGTGGTCACATTATCACGACGACTGCGCTCAATAACCAACTTTTTCCTCGCCAACTTGGCGTTTGCCGACTTCTGTGTTGGACTCTTTTGTGTGATGCAGAA
TCTGTCCATCTATCTGATAGAAAGGTGAGTGACAATTGGCAAAAACAACATTATTTCAGTATTAAATTAACGTTTCATCATGCACGAGCGGAGCTGCAAAATTCTCGGAAAG
CATATACGAAGAATACGTGAACTAACTTCACTGTAAAGTTGCCAATAGTTTTTAATGTCCGTTTTCATACGCAAAAAAATTTAAGTGAAGCATTTAAGTGAAGCATTTATAGA
AAAGTTTTTAGTTATAGTAGGTTAAAAATTAACACAAAAGTTAAAGTAGCGTTTCAGCTAGTACAGATAGTAGAATGTTATGTGGTCGATTAAATTTATTTATTTATCAGCT
GCAGAATGTGGCTGAAACGGTATCTGAGTGCATACGTATGTATGCATGTGCCTCGAAAACAAAGCCCCAAAAAAAAGTTTCAACTTTCAACCGTGAAACATTTACGCCGGCT
GTCACATAAAGGAAACTAAAACTCCGCAGGCGTCACTTTCACTTCAGGCTTTTCAGGCTTTTATCCCCACGCTCCGCTTTCACTCCAATCATGTTTATTTGCATTGTGTTTC
CTTAACTTTCCCAAAAACACGTCGCAGTGACGTGTTCACTTAAAATAAAGCGTGGAGTGCTGCTTGAAAATTCAAACAAAAAGAAGCCTTTTTCAGATTTGCAGGGCAAATT
AAAAAAGTATTTTGACAAATGAGTTTGGCTTACAGTTTCTATAAGTTATTTAATTTAAAATTTTAGTCGCAGTGTCAGAAATAAATAATGCTTCATTAATTTAAATTTGTGA
AATGCATGTGTTTAATATGTGTTTAATTAAAAGCTATGTTCTATACATTTAGAAAAAAAAGTAATTACCTATCTCATTAGAATTCTTGAAATGACCATTTATTTATGCACTT
AATAATAAACAGACATAATGATTTATTTGCATTCCCGATGTTAAATCCTGAAATTGGTTTGCATATTTTCCATTTCATTTTCTAATATGCTTTTTGTGGCCAAAGAACTAAA
TTATTGAGATGACCCATCTGTGTTTGGCATAATTTTAGCTAAATTATGCGCACAACTTGACGCCCACCAGAAATCATATGCAAAATGAGATGAGATTACGAAAAATCATAAT
TTATGCATTTATAATTTATATGTGCGGCATACGGCTGCCAGTGACAATTTAATTTAATGATTATAAACATCAAGTCGTTTTAAGAAAGTTTTTAGTTCATTATGTTAAGGAA
TGATTTGTCCCTTTTAAAGAGAGAAATTGAATCACCAGGCAAATGGGGAATCTTTTTAAGAAACGCATTATCCAAAATCTTGGCGACAACCTTCAGGATTATTATGTTTATT
ATAGTATTATTGGGCACATTTAAACGCATGAGTGCGAAATTACGTGGCATTGCTAATAACGAGGCACAAATGCGGTTTTCGAATGTAATTTCAGCCGAAATAGTAATAAAAA
GGAAAAAGCTTTGCGCCTCATAAAATACTAAAGACACCTTAATGTCGCATTAGTTTTGGCGGGCGTGTCATAACATTTTATGTGGCTTGTGCTAAATTTGGCTCTGCCACCA
ATATTGCCAAGTATATGCTCTTTAAAGAGCGAAACTTTTGCCCATAAATAATTATTTATGTCCTTTCAGCTGGGTCTTTGGCGAGTTCCTCTGCCGCATGTACCAGTTTGTC
CACTCCCTGAGCTACACGGCCTCCATTTTCATCCTGGTTGTCATCTGTATGGAGCGCTACTTTGCCATCGTACACCCCATTACCTGCAAACAGATCCTAACTGCGGCTCGTC
TGAGGGTAAGTTCTAACCATAAACCTCATATTCACGGCCAGTAATAGGCAAGTATGCATTTTAAATGAGGTCTATGTGTATGTGCCCTGCTTTTTTCAATTCCCATTTTGAT
GGGTCAAAAGCGGAGGAAATCTGTTTATGTATTGTATTTTGAGACATTCTCGCAACTTTAATAACTTCATTTGAATACAATTTGCAAGCTAATGAAACTTATTAGACCAAGC
CCGGAGACTCATGCAAATTAACCGGCAGAACAGAACCGACTCGAAAGGCTTAATAAGCCGTGGGCATAACTTTGTGGCCGGGAAAACTTTTCCGCACTCGAAGGGTGAAGTG
GAGGGAGCGCAGCCAAAGGGTTGAACCACATAGCCCGGTTGTTTCATCACAGCCAGACAAAAAATACGTTTCGGTAATGACGATTGCACTAATTAAACTACAAGAGACATAA
AGACAACGAAATGGGCTATCCATTAACATGCAAATTCAGTGTGTTCATTCGCAATTTAACAAATTGTTTGCGTGTTTGGCAGCCGAGCAGACAAAACAAAACGTCAATGCGT
GTTTTTTTTTTTGCAGATGGTCATTGTCACAGTTTGGATTACATCGGCAGTTTACTCCACGCCCAAGTTTGTTTTCAGCAAGACCATTAAGAACATCCACACGCAGGACGGG
CAGGAGGAGGAGATATGTGTCCTGGACCGTGAGATGTTCAACTCCAAGTTGCTGGACATGATTAACTTTGTGCTTCTCTACGTTATGCCGCTCTTGGTGATGACGGTAAGTG
TAACAAATGATACGTAGCTTGAAGTTAATAATTATGTTTAACCAATCCTTGGATAGGATGTTCATAAATTCATTTTCAATCCAAAAGTGTTTCCTAGTCACGTTACCTAGCA
TTGAGCAGACTATAATTAATAATTAAATGTATTTTTTTGAATTGCTTATTAAAGTGATTAATTATATCAATAACAACTATCTCTCTAATGAGATAGAATATAAATAATTTTG
GGTACATAATAAATAATCTATGAAATGATTTTTCTTCGTCTGTTAAAAAAGTGATTAATGGAAGTGTAGATTTCCCTTTATTTTTTTAACGACTCCATTTTTCAATTTCTTC
TGAGAAATTTGATAAGAATTGAGAAGCTTTTTGCATAAAATCGCCCCTTGTTCCGCAATGAAATTTGAATAATGAATTTAACGAACCCATTTGCATTGCAGGTGCTGTACAG
CAAAATCGCCATCGCCTTGTGGCGCAGTTCGCGCGGCCTCACGCCGCATGTGGTGCAGCATCAGCATCAGCAGCCGCAGCAGCCATCCTGCCAGGACATTGGCATGGGTATG
CACAACAGCATGTACCATCACCATCCGCATCATCATCACCACCACCACCAGCATCACCAGCTGCAATCGGCGGCTTCGTCGGCAGGAGTGGTGGGAGTGGGTTTGGGTGGTG
GTGGTGGGGGCGGTCCTGGTCCGTCACTCGCTTCCGGCGGCAGCAGCACCACGTCCTTGTCCCGCAAACAGAGCAGTAAATATGAGAAGCGCGGCGTCAGCATCACGGAGAG
CCAGGTGAGCTGCCAGGAGAACACTGCATCAGAAAAATCGCCGGTTGAGCACGGAGGAGTAGCCTTAAAAGCCTAACAATAGCAATAAACCGAAAGTGGAGTCGAACCTTTG
TTCCCTAATTGGCCAGAACGGCTTTAACTTCGATTGAGTCCGCACGTGCTGTGTGGTAAATAATGTTTTTTGAAAATCACATCGCTATCGGAGCACTAATTTATTAGATACA
CCATTTGTCGGCCTGGCCAAGTTGGCAGTCGTAAAAACTTTACCCGTAGCTGTTACTCCCGCACTTTCGCCGGGCCAAGTGGCACGGTGCTCGGAGCAGGTGTCAATTCACA
TTAAGGCCATCGCAGGACCATGCCACAAGCTCGGCCACCGATGTGACACCTGGAGTGTGGAGTGGTGGAGCGAAGTGACCACCGGCACAGGGCGAAACTGCAAGTGCGAGGC
CAAAGTTATGGCAAAAGTGTTTTTTTTTTTCCATTATCGGTGGCCGGACAGCAGGTTCAATTGGTCGGGTGCACCATGGCGTATTTCTTATTTCTTATTTCGAGCTCGACAG
CAGCTTTGCCACACAAACTGGCACACACATCATTGTTCTTGATCAGCATCAGCAGCTCACTTTGTCAATAAAAGCCGGCGGATAAAAATCGAGAATTTGTGACACGATGACA
TTCGGAAAAGTTGTGCAACTGCCGCAGATTGAAAATTGGTAAACACCAAAGGCAAGGCCCTTGAATTTTTTAATATTCCTCGAAGTATCAAATGTCAGGGGAGCACTGGTTC
GATTTCGGTATGTTTTCCGTTTTAAAAGTCACAAAAATGGAAGTTTGGTAGGAAAAACAATGGCGGGAACACCTTTATAAAACTTTTACATAACTTTAGTTTACGCTGTTCC
TTAACTTATGACCCCGATATACACAGATAACAACCTAACTATCTTTCTGTTGAACTCTTTTCTCTCTCTCTACCTTTCGTTTGATCAATAAAACTCTGTGCTCTCCCCCTCT
ACTAACTCCCCAATTATAACAACAACCACACACAATCCCACACCACAAATTAGCTCGATAATTGCAAGGTAGGTTGTAATCGAATATCAAAATCGATATGATAATGTTCAAC
AACTGTTGCGGCCAGCGACCTAATGCCATTACACCTGTTAATCCGCACAGGTGTCCCTGGAGGCCGATCGTCCCATTGTCAGTGCCTGCCGCAAGACGAGCTTCTACCACCA
CGGCCACGCCCATCACCAGCGGGCGGGCAACGCCTCAGTGGGCGGTGGCAGCGGCGGAGCCGGGGCCGGAGCCACCCACATGTCCCACTCGTCCAGCAACGTCCTACGCGCC
CGGCGCGGCGTTGTCCGCATGCTGATAATATTCGTGCTGACATTCGCCCTCTGCAATCTGCCGTATCATGCCCGCAAAATGTGGCAGTACTGGTGAGTAGTCCCTTCGTCCA
CCCACCACCCATAATCACAATCCAATGCGAATCGGAATCCCAATCCCTATCCCTATCCCTTACCGCAGGCCATTCATCAGCCCACACGGACTGAGTGTAAACAATATTTTAT
GGTATATTGCACCGGCAATCGGCGGGAATTTGCATAGTCCTCCGTTGGACGGGAAATTAACATCTGCGAGTGGGCAGATAAGCTCTAGATGGCTTCCTGCTGGAATCGAGGC
TAGTAAACAATTAAGTCATCAATGGACAGAGCTGGAGCTGCAGCCAGCTTTGCCGCTGGCCTAAATGATTTGAAAGCGGTCTTCCGTTTTCCAATCCCACTCCCTTCCCAAA
GACGCCAGTTTCGCTTTCTTCTCGCCTTTCAAAAGCGGTCGCCAAGAAAAGTCCGTTTTAGCCAGCCCTTTGGCTATCAGTTCCGCTTCCGGTGGAGCACTTTCGAATAGGT
GGACGGACCGCTTAGTTTGCTTTGAATTCGGGCAATTTGGTTTTTCCGGCTGGCGGGTGTGGTGGGCGTATAACTTGATAAGTGAGTGGGCGGTGGAAGTGTTTTTGTTTGC
TAAATCGAGGAGGAGTTAATTGCTCCCTGGGTAGAATTAAAAACGTGCACTTAAGCTGCTGACTGCTGTAAATTGGAAATAAAATAAAAGTCACAGAAATGTCCTTTAAGCG
AAAGTCGAAAGACAAACGATTGATTGAAGATGGAAAAGTCATCAAGTGTGAAAATCTGGAGACATTGGCAATTTTGGCACGTAAGAGAAACGTTTTTATTTTGAACTTTTC
ATACAGATTCTACTAAACCAAACGTTTTGAGTCAATTGTAAATTGTAAAGCATTTTACATGCCGATGTGCAATAAATCCAATTTAATGTTCATAAGTTCATGTTTTTGGTG
AATAACTACTGAAATAAATATCGATCTCTAGTTCTCATTAGTGCAAATATTCAGTGTAATTTTGTAAGAAAATACTTCGCTGCAGATTTGCAACGTTGCACTTTTGGCAAAC
GCAACGAAAACGAAAATGTTTTTGAATTTTTGTTCGAGCACTCAATTTGCACAATTGTAAAACGAATGGCGGACAAATAAATATAAATAATATCTGGCCAAATACAAGAGAA
ATGGATTTCCACATCGCCACTTTTCCCGCAGTGCCATAATTGAGATTTATGTCTAGCCCGGAATGGGCAGCGAGACTTGTTTACTTGTTTGAACTTTCAGCTTTTATTTATT
TGAAGAATTGGTTGCCCCCAGGGATATTGGTATTGTTAGGTGGAATTGAAAGTGTT
(SEQ ID NO: 172)

Exon: 1001..1032

FIGURE SHEET 75

Exon: 2646..2805
Exon: 3377..3577
Exon: 4022..4372
Exon: 4614..4820
Exon: 5539..5804
Exon: 5893..6112
Start ATG: 1001

Transcript No. : CT33559
ATGCAGAATCTGTCCATCTATCTGATAGAAAGCTGGGTCTTTGGCGAGTTCCTCTGCCGCATGTACCAGTTTGTCCACTCCCTGAGCTACACGGCCTCCATTTTCATCCTGG
TTGTCATCTGTATGGAGCGCTACTTTGCCATCGTACACCCCATTACCTGCAAACAGATCCTAACTGCGGCTCGTCTGAGGATGGTCATTGTCACAGTTTGGATTACATCGGC
AGTTTACTCCACGCCCAAGTTTGTTTTCAGCAAGACCATTAAGAACATCCACACGCAGGACGGGCAGGAGGAGGAGATATGTGTCCTGGACCGTGAGATGTTCAACTCCAAG
TTGCTGGACATGATTAACTTTGTGCTTCTCTACGTTATGCCGCTCTTGGTGATGACGGTGCTGTACAGCAAAATCGCCATCGCCTTGTGGCGCAGTTCGCGCGGCCTCACGC
CGCATGTGGTGCAGCATCAGCATCAGCAGCCGCAGCAGCCATCCTGCCAGGACATTGGCATGGGTATGCACAACAGCATGTACCATCACCATCCGCATCATCATCACCACCA
CCACCAGCATCACCAGCTGCAATCGGCGGCTTCGTCGGCAGGAGTGGTGGGAGTGGGTTTGGGTGGTGGTGGTGGGGGCGGTCCTGGTCCGTCACTCGCTTCCGGCGGCAGC
AGCACCACGTCCTTGTCCCGCAAACAGAGCAGTAAATATGAGAAGCGCGGCGTCAGCATCACGGAGAGCCAGTTGGCAGTCGTAAAAACTTTACCCGTAGCTGTTACTCCCG
CACTTTCGCCGGGCCAAGTGGCACGGTGCTCGGAGCAGGTGTCAATTCACATTAAGGCCATCGCAGGACCATGCCACAAGCTCGGCCACCGATGTGACACCTGGAGTGTGGA
GTGGTGGAGCGAAGTGACCACCGGCACAGGGCGAAACTGCAAGTGCGAGGCCAAAGTGTCCCTGGAGGCCGATCGTCCCATTGTCAGTGCCTGCCGCAAGACGAGCTTCTAC
CACCACGGCCACGCCCATCACCAGCGGGCGGGCAACGCCTCAGTGGGCGGTGGCAGCGGCGGAGCCGGGGCCGGAGCCACCCACATGTCCCACTCGTCCAGCAACGTCCTAC
GCGCCCGGCGCGGCGTTGTCCGCATGCTGATAATATTCGTGCTGACATTCGCCCTCTGCAATCTGCCGTATCATGCCCGCAAAATGTGGCAGTACTGGCCATTCATCAGCCC
ACACGGACTGAGTGTAAACAATATTTTATGGTATATTGCACCGGCAATCGGCGGGAATTTGCATAGTCCTCCGTTGGACGGGAAATTAACATCTGCGAGTGGGCAGATAAGC
TCTAGATGGCTTCCTGCTGGAATCGAGGCTAGTAAACAATTAAGTCATCAATGGACAGAGCTGGAGCTGCAGCCAGCTTTGCCGCTGGCCTAA
(SEQ ID NO: 173)

Start ATG: 1

MQNLSIYLIESWVFGEFLCRMYQFVHSLSYTASIFILVVICMERYFAIVHPITCKQILTAARLRMVIVTVWITSAVYSTPKFVFSKTIKNIHTQDGQEEEICVLDREMFNSK
LLDMINFVLLYVMPLLVMTVLYSKIAIALWRSSRGLTPHVVQHQHQQPQQPSCQDIGMGMHNSMYHHHPHHHHHHHQHHQLQSAASSAGVVGVGLGGGGGGPGPSLASGGS
STTSLSRKQSSKYEKRGVSITESQLAVVKTLPVAVTPALSPGQVARCSEQVSIHIKAIAGPCHKLGHRCDTWSVEWWSEVTTGTGRNCKCEAKVSLEADRPIVSACRKTSFY
HHGHAHHQRAGNASVGGGSGGAGAGATHMSHSSSNVLRARRGVVRMLIIFVLTFALCNLPYHARKMWQYWPFISPHGLSVNNILWYIAPAIGGNLHSPPLDGKLTSASGQIS
SRWLPAGIEASKQLSHQWTELELQPALPLA*
(SEQ ID NO:.174)

Name: Neuropeptide receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013385217
CCTTCGACTTGAGCGGACAGTTGAGGTTCTCGAACTAATTGCAAAATTGTCAATGATTTATTCAGTTGGCAATTAAGGTGCCAACTAACGGTGTGGCTGGCACGGCTATATC
GATTTGAAAATATTCAATAATCCCGGAGCCTTGATAAGCTCATCAAACATGCACATGGCCATCAGCAGGGGGGGGCAAGTTTGTATTTAGAGCCAAGTGTGTTTGGTGAAGG
GGGGTTAGCTGTATATTTCAGGGTTTCCCCTTTTTTTGTCGAGCGGTTTGGGGGGGGGTTTTGCGAGGGGACACAAAAGTACAAGTTGCTATGGATCAAGCTCAAGGCGTGG
CAGTTGGGCATTAGCTCTATGGTGGAGCACGGCAACCAAGCGAACAAAGGTCAACTCGACAAGTGAGTTAAATGACAAGGACCTCACCTTGCCTTCCTTGGGCTTCGGCTTT
GGCTTTGGCCCGTGCGTGTTTGCATGTTTGGCTCGATGAGGCGACGCGACCCCAAAAATGAAACTCAACTTTGGCTTTGACTGCGGCTCGGGATTGTGGCGTGCAAGCATTA
TGGCCTTTCGCTTTCAAAACAATGCCCCCATCCTGTTTCCCGGTTGGCCTAGTGGCACTTGCTCCTTGTTGTTGGTTTTTCGCATTGCCCTTGGCCAATTGCGAGTGAATTT
TGATGTTTTGACTGGCCATTCCTGGGGCCATTTGACAACTCTATAAAACCGGTGTCCTTAAGCCCGCCCTCGTCCTTGCGCCCAACCCCCAATTACCGTCCAACTACCACCC
ACCTACACCCCGCTACACCCACCCACAAAGTAAACGTAATTGCCCGGAGAGTTTCGCGACTTTCGACTTGGCCAGCAAACTGAAAAAGCGAAACTTCAAAAAACCCGACTTG
ACTTGGCAACTGGAAAAGTGGAAAAGCGGAAGCACGAGCTTTCCCACCGAAAAGCTCTCGTGCTCTAAGCTTTCGTCCTGGCGAGGCGAAAACGTAAATGTCAGTCATGGAG
TTCCAGTCGTCGGATACATGTGGAACACAACGAACACAGCAGGAGCCTCTCGCGTCTCATCCCCCGAATACAATGCGAAAAAGTCTGTACAAAAAATAGACAAATAAGGAGC
AGGAGCGAAAACACCACTCTGTGCATAAGATATAACTCCCGCAGAACTGAATACCTGCAAAAAGAAATACGCGTTAAATGGAATTACGGTGCGGGAATTGATAAAGTGAAGT
GTGTTACAGCTAGAAAGCAAGTGAAATAAAACGAATACATTAATAATTCATGCCCCTCATAGTCAGAAGCTTACTAGTTTGATAAGGAATTCAACGTAAATCTGATATAGTA
ATGTAGAAAAGTTGTACAAAGTAATATACATAGCATATAGCATATAGCACAGTTACACAGAACACACCGCAAACTCAAACAAAGTTGACAAAAAAAAAGCTCACGCCCCTCA
CATCTCAGTTTTTTCTCGAATTATTTGTGAAAGTGCAACAAACATATATAAAAAATAAAAAAGAAAAGGGAGGAGAGAGCACAGCAAAAAAAGGCAGTGACAGCAGGACCTG
CAAAGGAATTGCCATCATTTGGGGATTACGGAATTATCGGGATCCGTTCTCATTAAAATGGATTTAAGGCAATGCCGATGTGCTGAGCTGCATTTGTGCATGGGATTATCGA
TTGTAATGCGCAGTCATCAAATTATGTTTGAACAACGCCGCCGCGACAGCGATAATAGCAACAATAACAACAACAACAACAACCAAATCAACAACAAACACTGATTTGGCCC
CTGACCCCCTTTGGCAGCACCCCCTCCATCCAACACCCCCTACGACCCTGCCCAAGCAAAAGGAGCAGCGAAAAAGAAGAAGCGTCTTATGAAACAATTGAAGCACCGGCCA
AAGGAGCAAAGGAGCAGACAAAGACAGAGACAAAAACATCGGCAAAGGATAAAACGGAGGCCGAGAAGAAGAGGAAGCAGTAGTAGCAAAGCAGGAAACAACAACAACAACA
ACCGCAACATCGCATTTTTAAACGGCAATACGCAAAGTAAGTGAAGTGTAAAGTGAGCGAGATGGCAATAGAAAGGTGTAGAGTGCCGCGAAGAGAGGCGGAGAGAATACAC
TTGCAAAAATGTGAGCAGCTAATTCCCAGGGATGTTAAAAAAAATTAAACCCTCATGGTTCAAAGATTCTAGGAGATGGACAAGTGATGAATGGTCAGTGGAAATATATGTA
TGTTTTAAGCCAAGTGAAATCAAAATGATTTGAGAAGTACAAATAAAAGAAAGAAAGGAAAGAATGATATATGATATTTATTGGTTGTAGTACAGGAAAATAATAGTTAAAA
TGATATCCCGGATTGATGGTCATTACAATGATCGTACCACAAAACTCTTTCTAACTCAAATTAAATTGTTAAAATTATTTTTCAATACAAATTTGATATATTTTTGTTTTTT
CGCCGACAAGATGCATAGATCTTTCATAGAATCTGAAATCTATATCTTTTTTATATTGCTTTTGAGAAGAGTAAACCGATCATCATCTACGATCACTCCTTTTTGTTCTGTG
TAGGTATGAAGAGAGCAGGTGGGAAGGCAGTGAAGTGAAGTGAAGCATTTGTCGCATTTGTGCAGAATTCCCGACGTAATTAATGTTTAAAGTCTGTGTGTGTGTATGTGTG
CGAAGGTTTGTAAACTAGGTCACTAGGAAGTTAGCCCCTTGACGACCATCCCCTCTCATCAGCGCCTCTTCAGCGGCAACTCTACGTCAAAGGTGTCAACTCGCCAAATCTG
CAACCACACCTTAGTTAAGTGAGCCACCCCGTAGCCACCAAAATTATCCACAGCACCACCACCTTTTATTCGTCCTTAACACACCGCCATGATTAACGCGCCTCAGTGCGAT
TAACTGGCAAACAAATTTGGAGTGACGCTGCGGCGATGTACATACATACATATATGTATGCATGTTTGTGCTAAACCGACAGCATTTCCATTCAATAACACACGCGGCTCGT
ATGCAAACACTGTAAAAAGGTTTTATGCTAAAAAAGACGCCGTATTTGTGCATGAAAAACTGTCATTCATCATCCAACCAACACAAATAACCGGTAATGGCATCGCCACATT
ATGGCTTTGTATTTTTTACATCCTCAAACACAAAAAAACTTTAGTTGATTTCTGACTTGCATCGCATTCATCACCCACTAAAGTCGAGCAATCAATCCTCAAAATACGTATA
CGCACTGTTAGTTCAGTTGATTAATAGGCTAATTATGAAGGAAACTGAAAGGAAACACTAAATCAATAAATCCTGGCAAATTAGCATCAGTATATATTTGTGTTAGTGCAAA
CATTTAAGTTTTCTATCAATTTTAAATACAAATATGTTCAGAATTAAGATATATTTCAAAGGGAACTATGCTTTGTTTAATTAAGATTTAAGGTTTTATAACTATAAAAGCC
TTCAAAATGTGCCATCTATTTTATATTCACTTACAAGCAGATTTCAATCAAGACCACTCTTTAATATGATATATTATGTCTAGAACAACCGAGAAGGCACTCTTAAATTAAA
TATTTTAGTACCGTAATAAATTATGAAACGTGCTGACGGTAAAATATATAATACAATTCATACTTTAAGTTATATTTTTTTTTTTGCCCTATATGGTTTTACGTATCGGTGTG
TATGTGAGTACCTCGCACACGCAGGACATAGAGATAAAAAGGCTTCTCCCACATAAAAGTTCTATTCTGTTTTTTTTTTTTTATAGAAAAATATTGCTGGCAGGAGTTTAGA
ACATTGCCAAAGGGGATGCGGACAGATGAAGGGGCGGGGGCGTGGCGGGTAAAAGCCATAAAACTCAGGCGCCGAAGCGCAAAATGCCCTTGCCCTGAACAAAAAGAAGAAA
TTGAACAGCAACAACCAACAACAAAAACAAAGAAGAACAAGAGCAAATAAAATACAAACTGCTTATTTATTTATTTGTTTCGACAAATCTCAATTGAAACAAAGCTTAAAAG
CTTCGAAATCAAAACTTAAATAGAAACCACACAGACATCCATCCAGACGAATAAAACGAAAACGATAGAGCGCGAGACAAAGAGACAATTACTATTAAATCCAAGTTCAATT
CGGTCCGCAAACAATTGAATTTGTTGCAGATTTGATTTGGCTTCATGTGGATTGCGATAACTATTTTCCTGGGATTAGTATGGGATTGGATTCATTGGCCCGACCGCAGTTT

GCTTCCTAATGAGATTTGCTTAATTAAAATGAAATTCAATAACTGGAATTCCAATGAATTTAAATGTTAGAGAAAAACTCTAATTCGGACTTTTGTGGCTAGCCATTATTTG
CAGCAAGGGGGACATAAAAAGCATTTGGCTTTCGTTTGGATTAGTCAGCAAGGTCAAATCTCGAAACTGACCTAAGCTAATGCAAAATGACTTTTCCTGTCAATCACAATCT
ATATTTAATTCAAGCGAAATTCTCTGCCACATTTCTTGCGCTATAAATCGCAGTACCCATTATAGCATATCAACAACCAAATTAAGACGAAAGTACAATAAAATGCAACTAA
TTTCAGTTTTAATAATGGATTTGCTGGCGGCTTATTTTTGTTTCAAATTTGGTTATTTTCTTTAGAAATGTGCGCATACATTAAGACGAACACAAAAACCAGCTTTGGCTCA
ATCTCTTAAACCTTTAAAACGTTATATGTACGCACATATTTATGTACGAACAAATGAAGGAAAATCTCAAGCCTAATTCCGCATTAAGCTGCAAGGAAACACATAACCCAAA
AGTATGCAACATAAATAAAATTTAAATCAGCACGTTTATGAAGTGGATTATTTGAACTGAAATGCTTATTCAACGTTATTAGATGAGATTTATGGCCAGGAGAGCATTTGAA
TTGGCAGATTATTCGGGCCAACCAAGTAATGGCCACTATATCAAACAAATCAGTTATCCTTCAATTGCAAAATACACAGTGCAGACAAATAGTTTAAACTGTCTTAGTGGTA
ATCAAAATCAACTTTAATTATCAACTAATGTATGTGTATTTTACGAAACAAAAGTGCATTATGCATGAACGAATGATAGCAATATTTTTCAATATAAAAGCTTACGATTGTT
TTCCTTTTGTTATCAATTATCTATCACACTCCACAGACATACTCCTAGATCATAATAATAAGTTATATAAATATAGCATAGATTTTGAAAGGTTAAAAGGTTTTCTCCCCAC
CACTCCCTCTAATAATATTGTATATAATAAAATATATATACAAATGTATTTCATAATTTCATTGCCATCCAATTCCGCATAGTATCTGGTATACCCCATCTGCAGAAAGACA
CCGAAGTGTATAAATTGCTTACGTTTGCCCATGTAGTCATAAAAATTGCTTTGGCACATAAAACGAAGCTTGACGCTCGCTTTTGGTTTTGGTTTTGGTTTTGGCCTGGTCT
AGTCTGGTTTGGTTTCATTTGGCCCCTTACCGTTCTTGTTTCCTTTGGTTTCCATCGCTGGTGGGCTTGTTTTTTTGCGAGAGTCCTGGCCGTGTCCTTCTTATATTTCTAT
TTCTATTTGTATTTGTATTTGTATTTGTATTCCCCGATTTGTATTTTGTTCGCTCGCCAATTCGGTAATCACATTTCAAGCGTTTTATTCGTCCTGATTTCATACGATTGCT
GATTCAAGTTTCTGTTTGTGGTCCTTGCCTTGCAGGATATGACAAGTGATGGTGCTGTTACGTTTTGGATATTTTTGCTTTGTTTAATCGCCTCGCCGCACCTGCAAGGGGG
CGTGGCCGGGAGGCCCGATGAACTGCACATCGGCGGCATCTTTCCGATAGCCGGCAAAGGAGGATGGCAGGGCGGCCCAGGCGTGTATGCCTGCCGCAAGACTGGCGTTGGAT
GATGTCAACAAGCAGCCAAATCTGCTGCCGGGCTTCAAGCTCATCCTGCACAGCAACGACAGTGAGGTGAGTAGTCAGTACACTGAAAAAAAAACGATATATACCTAAGGAT
ATTTGAATGTAATGCGATGCAATAGTCCTACAGTCCAAGTTCGATACGAAAATAGTTTTTTATGTTTCCTCTTAAGCTGGCTTTAATAAGAATAAACATTTATTGAAGCTT
TTGAACTTTCTTATAATTTTCATAGGTAACTAATATATACAATATATTACATTAAATACCCAATAAGTATCCAATAAACTTCGAAGCAACACAGAACTTGTTTGCAACACGA
TTAATATCACTGCAAATAATATCCTGACCAAAAACCCCTCCTGATGCACTGAGTTCTGTTTAAGCCTTCTGTTCACGAAGCATTTTCGCTCAGTGCACTAGTAAGTTGGAAA
GCGATAAGCCCTACTTATATGCAGCCCCTGCTGCCTCCTCTGCTGCCATGCCACATACCCTTGGGTGCCACTTATTTTGTTTCGTTTTTGTTTCTGCTCACTGATGAGTAGT
GGAGTGCTGGCTGAGCGCTGGCATTCATTAGCATATTCTCTAGGGGATCAGGCTGGGGCTGGGGCTGGAGCTGGGCACTGTTAACAAAGGACAAACAGTTTTATTATGCCG
CATTAAAGGGTATTGCTGCGCGTGTGTTCCACAAATTGAATTCGCCGCACAATCCGCCGCATGTCCAACAATATTTGGCAATGCCTCGGCTTAATGACTGCACATAAATTGC
GATTGAGATTCGACTGGAGCACGTTCGAAACGTTACCAACAGCAATTATATTGGCTGAGTTCGGTTAATCGGGCTTTTATGTGCGTGATTCGTAATTGATAAAGGGTATTTA
GCATTAAGAAAAATGTGTGTATATCTCGAACAAGGTCGCCCAATAAATGTCGCCCGGTGTTCATTAAATCATAAATGTATTTACATTTATTCGTGTAGTGCAACGCTGTAAG
TTTAATTATTTCAATGCGAAAATCTCTCATTTCCCAGAAAGTTTCTATGAGATCAATAAGTAAAAGTTTTTGCTATAAACTTCGAACTTGGTTTTTTGCAATCAAAAGTTTT
ACGCATTTTATCTTTAATGCCAGCAGCAGCACGAACTGAACTTCTGCACAAAGCTTTTCTTGAATAAATCGGTATAATCAAAGTGAACACAGTATTGATCATATTTTTTTTA
TTGATCCTTATATAAAATTGTGAAATTTTTGCATTTTGTAGCTAACTTAAAGAGCGCGGAATATGTTTTTAGCTGGCATATTTGTTCTAAATGGAAAAAATAGAGCATGGGC
ACGCAATAAATCAACATTTCAATAGTATCCAGTTGTCTGGATGCTGTAACTTCTGTTGGCTTTTAGAGAACCAGCTATTAAACAAATACCCGCCCTCCAAACGGTCATGCAT
CATAATAAATGGCCAGTAACCGATGTAGTTTCCGTTTTCCAGTTCCCAACCAAGTTGCGCGCTATCAATGCCCCTTGGCGCGTATCTCTAATACAATACATATAAAACCAAA
ACCAAAACTATCTTTATACATTTTCATATTTCGCACATATTTTTTGTTTTTTTGTGAATTTCGCATTTTAAGGGTCTGCGGTACAGTTGTTGCGGTTTCCACTACTTCGGCG
GCTGTTTTGGGCCTGCCTATTTTTAAGCGCGAACCTCATTAGTTTTAAGTGAGCGAAAGCTTAACTACTTGTTTGAGTTTTGAATGAGGTCACTTTGCCCCGGCCGAGCCG
TAGCTCATCATTTGAGTTATTTTTATTCAAGGCTGCCAGGAGGGACAGTCACTACTACTACATACATATATCCTACGAACATATACGAACATATACACGCCCATTTACGAGC
CAGATACGGACATATAGATAGGAACAGAGCCGGAGAGGTAAACCAAGTCCTATCCCTGGTCATATCTGTTCCTATCTGGCGTGGTCTGGTCTGGTCTGGTCTGGTTAGTTCA
GTTCAGTTCAGTTCGCTTCGCTTCGCTGGACTGAGCCTGAGAGCCAGGCGCGTTATGTTTGTGGCTTAGGCAAACAGCTTCCGTTAAATGCTGCATAATTAGTTGCCATCTC
CAGGTCCCACAGCCACCTTCTCCTCCGGCACATCCACATCCACACTTACTCCCACAACGCAACGCACCACCCATCCCGCCCACATCCACGTCACTTGGCAAGCGACAAGGG
GCAAAGGTCGCACGCAATGTTTGCCTTTGCCGCTGCCTTTTGCTTGACCTACTTGATGTGTGAAATGTCCCAAAGAATTCTGCACAAAAGCACTGAGAGTTGAACAAAAAA
ATGTATAAACAAGGCAGGTTGGTTATCCTGTAATAAGCGAAATCCCTAGTCGGCTTAAGTCCAAGATCACACAGAACAGGTGAATACTTTTGTGGCGAATTTCGAATCGAAC
TTTGATTTTTTAATGGCCAAATATAGGTGCTGAAAAATCGTATTATTGCATTTGCAAAGCTGCCACTTTCTATCGGCTAAGTGCCTTCATGCGAACCACTGAAATGGCTTAA
ACCGGTTGTCAAGCTTTAGGAATAAGCCGGAAATTGATTTAACAGCGTTGGGATAAAAAATATGTGCTCCTAAAAGGTTTTTGCTTTTTCTTTTTTAATTAAAGTTTAATTAA
ATCACAACAAAGCATTCAATTATTTAAACTCACATTTTCAATTGCAATACAGAAAAGGTAAATATAATGGCTAAGGTAATCAAAAATTATATAAATAGGCATCTACAGAATC
TGCGAGTTTTAATTTAAATTGAATATTGTTTTGTACAATGTAATAATATTCCTAAATGTGCTTTGTATTATAGAAGAACATTCTACTAATAAATCATTTAGCTATTTTTTGG
TTAATATTAAAAGTACTATAGTACTATGTACTATCAGTTTTATGCTGCTTAGTTGTTCTATAAAATCAGTATGCCTCTGCCCCACGAAAAGGGAAACTATTAAATTAGAGCGG
ACATCGTGGCATGCCTCGGTACCCAAAGCCAAAGCACACACCCAGATCTGAATAGATGGAGTAGTACGAACAGGATAATACTTATGCATAAACAGGGACTGGACATGAGAAG
GAAATGAAAGGAAAAAAAAAAACCAATTAGAATGGTTGTTTACACGCAAACCAAGTAAGAAATATGAAAATTGTCGTCAGAATGGGGTTTGGGCCAAAGTAAAGGAAACGCCG
AACAGGGAAACGAGTTTGCCATGTAAATGAATGAATGCGAGAATGAATAACTGAATGAATGAATCATTTATGTGGCCAATGTGGATGGATACACGTATGCCCACGTATGTAT
GTATGTATATTCACAACTTCGGCACAAGGGAAAGGGACACACCTTCGTGAATTATTCATTGTCATGAATGAATCTTGTTTAATTGATAAATGCTATGACATGGGGATTAATG
AAAAGCATAATGGAAAAAATATCAACCAGGTGGCACATAAATTAGGCAGGACTTGGAATATGGAAATTTAACTAGCTACATAAATATTGCTACAAATTCGTATAATAAATCAC
GCTAAATGCAACAATTAACATGCATAGTTTAAAATTTGATTGATTATAATTTTCAATTTTCTCTGCCAGCAAGTTTTTAATTAAGCAGCTTAATTTCGTAGTTTGCTATGTC
AAATCCAGCATTCCATCCAGTTGATTAATTTTCATTTTCAATTATATTGGCGAAAATTGGACGAGATGACGAAAATTCCTTCTGGGCCAGTTTTGTGCATTAAATTCAATTT
AAAAGCCATGCCAATGTCATAAACGCCCCGCATGAATCACAAAGCCAATTAATCATATCAATAATTATCTTCCAATGAAGACAGACACATAAATAACAGGAAATTCGCTATT
AATAACTCCTGAAGGGTCGTTGCTTAGCATAATCATAATCAATACATCGCGTTGGCCATCAAACGAAACCTACATATCCGCCCATGTTTATGAATATTTATCGAGCACCGAC
GGATGAGAGTATAAGTATTCTTGGCAACTTGGCTAATTCCGCTTTCCCCATCCCCATCCCACCATTTTCATTCACAGTGTGAGCCCGGTTTGGGCGCCAGCGTGATGTACAA
TCTGCTCTATAATAAACCGCAAAAGCTGATGCTGTTGGCAGGATGCAGCACGGTCTGCACCACTGTAGCCGAGGCTGCCAAAATGTGGAATCTAATTGTGGTAAGTGAAATT
TGCATAGCGCCAGAATGGCTGGATATATACATACATACTTCGAAATACTTTACCTACATGCACTGCCCCTCCACAGCTCTGCTACGGGGCCTCGAGTCCGGCTCTTTCGG
ATCGCAAACGATTCCCCACTCTATTCCGCACCCATCCATCGGCCACGGTGCACAATCCAACGCGCATCAAGCTGATGAAGAAATTCGGCTGGTCCCGGGTGGCCATTCTGCA
GCAGGCGGAGGAGGTCTTTATATCGGTAGGTAATCGGTGCTAACTAATTACATTTCCATATTAACTTCGAAAACTTTTATGAACAGACCGTAGAGGATCTCGAGAATCGATG
CATGGAGGCTGGCGTTGAAATCGTAACTAGACAATCATTTCTATCCGATCCAACAGACGCCGTGCGCAATTTGCGACGCCAGGATGCACGCATCATTGTGGGACTCTTCTAT
GTGGTGGCCGCCAGGAGGGTGCTCTGCGAAATGTACAAACAGCAGCTATATGGCCGAGCTCATGTGTGGTTCTTTATTGGTAAATACAAATTATACTACTTAGATACCGATA
TCTCATATGCATAGATACTTTCACAGAGTCTTCTTGTAACTAAACTACTTTTTTTTTAGGCTGGTACGAGGACAACTGGTACGAGGTGAATCTGAAAGCAGAGGGCATCACC
TGCACTGTTGAACAGATGCGAATAGCTGCCGAAGGACATCTGACAACGGAAGCGCTCATGTGGAATCAGAACAATCAGACAACTATATCCGGAATGACTGCAGAGGAATTTC
GGTTAGTACTACTACTCTATCCATTGCTAAAGTTTTTGCTTCAATGTGCCTTCTTTTTGCGCAGACATCGACTGAATCAGGCGCTAATCGAGGAGGGTTACGACATTAACCA
CGATCGCTATCCGGAGGGATATCAGGAGGCGCCACTCGCCTACGATGCAGTGTGGAGTGTGGCTTTGGCTTTCAACAAGACCATGGAACGATTGACAACCGGGAAGAAATCT
CTGAGGGATTTACCTATACGGACAAGGAGATTGCCGATGAAATCTACGCTGCCATGAACTCCACACAATTTCTGGGTGTATCGGTCAGTTTGTGAAACTAATTTCATAATA
CTTAGTTAGTTACTTATTTTTTTATTTGCAGGGTGTGGTGGCATTCAGTTCTCAGGGCGATCGTATTGCTCTTACACAGATCGAACAGATGATAGACGGCAAGTACGAGAAG
TTGGGTTACTACGATACTCAGTTGGATAACCTATCCTGGTTGAATACTGAACAGTGGATTGGTGGCAAGGTAAGCAAATCAGTTAGCGTTGAAAAAACAATCATATAATCAA
AATGTCGGACGATGATAAAAATCAAAAGAAGACAAAGCATGTAACGACTGAACCGAATAAGAACGGAAATCCGTCCAAAGACGACGTTAATACAAGAGGATCATTAACAAGC
TTAAGTGAGAAAAATGATGAATCCAAAAAATCCATCATCATCTGCAAGTCTTCAGCAAGCTCCTAAGGTAGCACCAAAACCGAAAAAGTTATCAATTTCGGATGCTGGTAAAG
ACACAGTAACCCAAAAAGTCAAAGAAAATGAAGAGCAGGATTTTCTAAACAATTTGAAAAAGAATCTATTGGTGTGCGAGCAAATAGAAACAGCAGTGCAACAAAAAAGGA
AAATGAAAAGCTATTAGTAAAGACCGTGCCCGGAAAATCTCTTATAAAAGAATCTAACGATGAAAATGTGGAACCATCTAGACGCACAAAAAGTCAGCCAGTTGGAAAAAAA
GTTTACGAGGAATCTTCAACGAGGCGAGTTCGTGAGCCATTTGATTCTTTTGATAGGGAAAAGTACCTGTCTGATATGATCGATTATGATCGATCGTCCGAGGACGTGCCGG
AAAAAAAACTATCTTGGGAACCAGACAGTACACTTCGACGTCGTTTTGTCAGCAATAATGAATATCATTCTAGTCTTGAAGGTGAGGAGGAAGAGATCGATTTAGATTCAGT
TGGATCGTCAACAATGCGAGGTAGCTATCGAATGCCACAGAATGCCGAGGATGAACGACCTGTTATCCTAGCTGAAATAATCAATATGGGAAAGTTGAAGGAGCTCAAGGAG
CAACGCGAATCAACTAACGTATTTCGTAACAATCGTAGAGATGATAAAACCATTGATGAAGTGAAGGGTTCCAAGATCACAGGAACCAGCAAACAACATGTTGAAGAAGTTA
AATATAATGAAACAGCAAGTAATGAAGATGATGACAATGCACAGGTAAAACATATAGGTACAATCACAAAAATCTTTAGCATTGCGCAAAGAATGAAAGAGAAGAAAAGCTC
CAAAGAAACTGAAGAAGACGAGAATAAGGACGAGAATAAAGATGTAAAAAAAGATGAAAATAAAGACGATGAACAGAAACTTCCCCCGGAACCAGAGATCGAATTGGAGGCC
AAAAAAGCTTGGACTTTTCCGATCGCAGAGACATCACCTGATATCCCTGGAGAAAATCAGATCCTGCAAGAATACCAAAAGGTTAAAATTGGCAAAAGAAACAGGAAATACA

FIGURE SHEET 77

GATGTATAGGCATCAATACGGATATAAGCAGAAAATCTAAGATTAAAACATTTATCAGCAAGTCTGCAGGTGATGACTTAGTAGTATACCATGATGGTGGAAGATTGCGAGA
TATAGGACAAACCACTGAAAATCTTAAGAATCAACGAAAGAATCATAAACACAAAACCGAACCAGATGATATACCCGTTGATATAGGACACAGCGATGATCGAGTAAGGGAA
ATTGGAGTGAATACTAAAAAGTTACCAAAAATCATAATACCACCAATTGCCGAGATGCATGTGCACAAAAATGGCAAATTAAGGGACATTGGAACCAGCACAGATAAGCCAT
TTTGGCCAATCGATGACGGAACTGATGTAATCTACATGCATCCCATTAAGACTGATAGGAAGAAATTAAACAAACTGATTGTGGATCCACCGCCGGATAATGGACCCTACAA
GATGCCAACCAAAGAAGATAGACGAACGTACTACAAAGGTTGTGAATATCATTTTCCTGGCCGAACGGAATGGCGCAGGCTTTTCTTCAACAAAATCCATGGAAAATACAAG
CTACGCCGTCCAAGTCATTGGCTCTACACACTAGTTTTTCCGTTTTATATATTTTGTTTGTCATTATCTTTTCCATGGCCTGGTTCGATTTCATCAAAGATGACGCTTCCA
GAAAGGTACCCATGATAAAGATGGCACAACCATTTATTAGTTTCACTCCCATCGGACCACGAACTAATCCCAAGGCGGTATCCTTTGATCCTAGAAACAGCACAGAAGTCAT
GGAAAAGTACGCTGGTATAATGGCTCTTCTGGAAAAATACGGCGACTATGGACACAATCCGCGTTTTGGAACGTGTACTGCGAATGAAAAGTTTGGCTATCCAAGCGGCGAA
CCGTGTGTTTTTCTTAAAGTTAATCGGATTATTGGATTTAAGACCGAACCCTACATCAATTCCGATGAGCTTGTCAAAGCCAAGATCGATGAGGTTGAATTTACGGCCTTAA
AAAGGTTATTGGAAAATACAACAACAGAGGAGGGTCATTTGAATCGCACTTGGATAACATGCCGTTCAGATAAAGATAAGAATGTTTTGATTGAATTTCATCCAGAGCCGGC
GATAAGAACGGAATACACCGATATTGAAGAGAAAATCGAGTATATTGCGAACGAGGGTAAAAAGTCATTTTTCGGCCCAAACGATGTGAATCGAATCGTGGCTCTTAAGATC
AAAAATCTCAAAGCAAATGAACGTGTCCATATAAATTGCAAAATGTGGGCTCAGAATATACATCATAGGAAAGAGGGTTATGGCCAAGTTTCGTTTTTTGTGTTATTAGCTA
CAAATGAAAACCGCGAGAGGGTGGAAAAATTACTCCGAACCGACTACTCATTATAAAATACATAAATATAACTTTCTAATAAAATGTGTTATATGATATATGTATGTATTAT
TTATAACATCTATAATCCAGGTTCCTCAAGATCGCACAATTGTCACCCATGTTCTACGCACCGTGTCCTTGCCATTATTTGTGTGCATGTGCACAATATCCAGTTGTGGCAT
ATTCGTTGCCTTCGCCTTGATCATCTTTAATATATGGAATAAGCATAGAAGGTAACTATTTAATTGGTAGTGCTCGTGCAGAACTAACTAAGGGTTATATAATTTCTAGAGT
AATACAATCCTCGCATCCCGTTTGCAATACGATCATGTTATTTGGTGTCATCATCTGTCTAATATCTGTCATCTTACTGGGCATCGACGGACGCTTTGTCAGCCCCGAGGAA
TATCCAAAGGTGCGAGTTACTTACCTTAAAAGTTCTCTTCATTTTTATTAAAGTTAATTTGATATAAAATTATCCATTAGATATGTCAAGCGCGGGCTTGGTTACTATCCAC
CGGTTTTACACTAGCATACGGTGCTATGTTCAGCAAGGTCTGGCGTGTGCATCGTTTTACAACAAAAGCAAAAACTGACCCAAAGGTGGGTTTAAATTAAATTAACTACAAC
TATATATTAGTTGCTAATCATTTTGTGGAATTGCAGAAAAAAGTGGAACCTTGGAAGCTATACACCATGGTTTCGGGGCTATTATCAATAGATTTAGTGATATTACTCTCAT
GGCAGATCTTTGATCCGCTGCAGCGTTATCTCGAAACATTCCCACTCGAAGATCCAGTATCTACTACTGATGATATTAAAATACGTCCAGAGCTTGAGCATTGTGAAAGTCA
ACGCAACTCCATGTGGTTGGGTATGTTGTGATAACGATCCATTATATTGTATATAATATATTTATATTATTTTGCAGGTCTTGTATACGGCTTCAAGGGGCTAATCCTGGTG
TTTGGCCTCTTTTTGGCGTACGAGACGCGCTCCATTAAAGTGAAACAGATCAACGATTCGCGTTATGTGGGCATGAGCATCTATAACGTGGTCGTCCTTTGCCTGATAACAG
CTCCGGTGGGCATGGTCATTGCATCGCAACAGGACGCGTCCTTTGCCTTCGTTGCTCTAGCTGTGATATTCTGTTGTTTCCTAAGCATGCTGCTGATATTTGTGCCAAAGGT
CAGCCACTACACTGAACTAAAAGAAGTAGCACATTTGTAATGATCTCTATTCGCACAGGTCATTGAGGTTATACGTCATCCCAAGGATAAGGCCGAATCGAAATACAATCCC
GATTCAGCCATATCGAAAGAGGACGAAGAACGCTATCAGAAACTTGTTACCGAAAACGAGGAATTGCAACGATTAATAACACAGGTATGCTGAATTATGGTAGATTATCTTA
ATAAAATCGTATTTTATTTAATTTTGTTTAAAGATTAAATTCAACTGTTCGTTGAACGTTAATCTAAAAGTATTTCGGTCTTATCATAATTTGAACAAAAGTTAAGGATTGG
AATTTCCAGCGGCCCCAGTGAATTCCATTCAGCATGCCGGTTTCACCTTCTCTAAAATAAAGTCCATTTAGCGAGCTGCAATATGAAGTATTATCAACCTCTTATAAATCAA
ATTTTAATTTGATAATTAATTTACCTTGACAAACATGAGTGAAACCACCAACCACCGCCATGCTCGGCTGCACAATTTTTACTAGATTCATCATTATCACGATCAAAAGTTG
AGAACCTTTTATTTATATGGTATCTTAGGGAGTCCCCAGCTGTTCCAGAATATTTTCCCACCCGTTCCAGTTTGTATAATTCAGTTTCACTGTCCACTTGAAAGTCATCAAA
GTGAGCGTATACTGTAGCGCCAGGACCATGCTTCAATTGAATGAAAAGTTCGTGTGGTTGTTCTCTAGTCATTAGGTAAAGCTTCTGAAGACCGATGAAAAATTCGCCATTC
GGATCGCCAAATCCATCTTTGTAACTAAACCACGATTGGTTAAAATTAACACTTCCATCTAGTCGTCTTTGAATGACAATCCAATCTCTTGCCGTCGTTTTGCACTGGGTCA
CTTGAAAAGGTTCTTCTTCTGGCAACATTAGCTGGTGAATTCCATTTGGACTGCCACTAGGACACGTGTGTGGTGAACTCTGGCCCACTTCGAACAGCAAAGTCCATAAAAA
TAGTACAAAAAAGCACGATTTCATTGTACTCGACATGCGTACTGTTAGAAGATCGGTGGCTTACTGAAGTATCTACCATTTAAAAGGGGAACATTGGGACGCAGTTTTGTGG
TTTTGCTACCTTTGACCGCATTGCTTGATATGCATATATTCTATTTAGGTCGGCATTTAAGCAAATGGATTCCAAAAAAGTGAATGGAGACTATTTAAAAGATAGAACTACA
TTTTTATTGTTTAAATGTTTGTTTGTTTTGCTTATATTTTTATTGTTTTATAAACATTGATTTATTAATTTTCCGCTTGCAGAAGGAGGAAAAGATTCGAGTCCTGCGACAG
CGTCTGGTGGAGCGGGGCGACGCCAAGGGCACAGAACTGAATGGTGCAACAGGTGTCGCCTCCGCCGCCGTTGCAACAACTTCGCAGCCCGCTTCCCTCATCAACTCATCAG
CACATGCCACGCCCGCAGCCACACTCGCAATCACACAAGGTGAGTAGAGTACCTGGTGCATGGAACTATGCATCCCAAATGCCCCTTAGCCAGTTGCATGTGGTAATCGATA
CTAGTTACGTTGGTTTAGCTTGACGTTCTTAATCGTTATGTTCTCGATTTTAACTAGAAATCTTAGCTCTATTCGTCGGCATTGACACAATTGATTAGGAACAACAATATAA
ATGGAGTTCAAAGGTTATTTGCTTATGCTAACGCTATGTAAACTCGATTTTCGCACAAAAAACCAAAGGCAATTTATCAGTTTCTGGAATGTCTATTGCTTTGTGTGCATAC
ATATGTATCTTAAGCACTTTTGCAGTTTTTGAGCTTATTTATATCTTCAGCACTTTAGCACTTTTTTAATAAATATCACTGCTTCTGCACCCTTTTTTCACATTTTAGCAAG
ATTATTCAAAACATACTAATACGTAATATATTTATCTTACAGATACTCATGAGCATCGTACATGAAAAAACTTGTCAATTTCCTTTGCGTGATAGGTGAATTATACAAAAGA
TATGAAATGCAAATAATGATGCTAACGGAAGGATGGATTTAAACAAACACGTACTTAAAGAACTTTGTATCAACTTTAAGCTGAGCTAAATAAGTGAATTGGTAGCAGCAAT
TTTGTAGAATAATGAAATATGTATTGTACCTTAAACGTGTTTAGCATTAATATTGAAAACGAAGTAATTTTAACTTAAAAGAAAACAAAACAAAAATTCGACACTAGCATCT
AAACTAAGCGAATTGTTAGGCTTTAAAGTTGCAACACAGAAACCAAAAAAAATTCTAGCAAATCAAGAGAATTCATTACGAAAAGATCGTGTTAAAGTTAGTAATTTTTAAA
TGCAGGTAGAGAAGACAAATTGTTGATTTTGTGTTTAGTAAATACTGTTTTTTTAGTTAAACGAATTATCCTTTCGAACCTCATTTGCTCAGCGCTGTCTCCTAGAGCTCTC
TAGTTAATTATATCTGCGTAGATGAAAATTCCCCAAAAG
(SEQ ID NO: 175)

Exon: 1001..1090
Exon: 1165..1208
Exon: 1810..2052
Exon: 3783..3856
Exon: 5748..6002
Exon: 10046..10180
Exon: 10271..10441
Exon: 10503..10719
Exon: 10812..11284
Exon: 11344..11493
Exon: 11565..14044
Exon: 14446..14569
Exon: 14902..15028
Exon: 15086..15342
Exon: 15403..15540
Exon: 16659..16847
Start ATG: 1026

Transcript No. : CT35221
TCATGGAGTTCCAGTCGTCGGATACATGTGGAACACAACGAACACAGCAGGAGCCTCTCGCGTCTCATCCCCCGAATACAATGCGAAAAAAACTGAATACCTGCAAAAAGAA
ATACGCGTTAAATGGAATTACGCACCCCCTCCATCCAACACCCCCTACGACCCTGCCCAAGCAAAAGGAGCAGCGAAAAAGAAGAAGCGTCTTATGAAACAATTGAAGCACC
GGCCAAAGGAGCAAAGGAGCAGACAAAGACAGAGACAAAAACATCGGCAAAGGATAAAACGGAGGCCGAGAAGAAGAGGAAGCAGTAGTAGCAAAGCAGGAAACAACAACAA
CAACAACCGCAACATCGCATTTTTAAACGGCAATACGCAAAAAAAATATTGCTGGCAGGAGTTTAGAACATTGCCAAAGGGGATGCGGACAGATGAAGGGGCGGGGGCGTGG
CGGGATATGACAAGTGATGGTGCTGTTACGTTTTGGATATTTTTGCTTTGTTTAATCGCCTCGCCGCACCTGCAAGGGGGCGTGGCCGGGAGGCCCGATGAACTGCACATCG
GCGGCATCTTTCCGATAGCCGGCAAAGGAGGATGGCAGGGCGGCCAGGCGTGTATGCCTGCCGCAAGACTGGCGTTGGATGATGTCAACAAGCAGCCAAATCTGCTGCCGGG
CTTCAAGCTCATCCTGCACAGCAACGACAGTGAGTGTGAGCCCGGTTTGGGCGCCAGCGTGATGTACAATCTGCTCTATAATAAACCGCAAAAGCTGATGCTGTTGGCAGGA
TGCAGCACGGTCTGCACCACTGTAGCCGAGGCTGCCAAAATGTGGAATCTAATTGTGCTCTGCTACGGGGCCTCGAGTCCGGCTCTTTCGGATCGCAAACGATTCCCCACTC
TATTCCGCACCCATCCATCGGCCACGGTGCACAATCCAACGCGCATCAAGCTGATGAAGAAATTCGGCTGGTCCCGGGTGGCCATTCTGCAGCAGGCGGAGGAGGTCTTTAT

ATCGACCGTAGAGGATCTCGAGAATCGATGCATGGAGGCTGGCGTTGAAATCGTAACTAGACAATCATTTCTATCCGATCCAACAGACGCCGTGCGCAATTTGCGACGCCAG
GATGCACGCATCATTGTGGGACTCTTCTATGTGGTGGCCGCCAGGAGGGTGCTCTGCGAAATGTACAAACAGCAGCTATATGGCCGAGCTCATGTGTGGTTCTTTATTGGCT
GGTACGAGGACAACTGGTACGAGGTGAATCTGAAAGCAGAGGGCATCACCTGCACTGTTGAACAGATGCGAATAGCTGCCGAAGGACATCTGACAACGGAAGCGCTCATGTG
GAATCAGAACAATCAGACAACTATATCCGGAATGACTGCAGAGGAATTTCGGTTAGTACTACTACTCTATCCATTGCTAAAGTTTTTGCTTCAATGTGCCTTCTTTTTGCGC
AGACATCGACTGAATCAGGCGCTAATCGAGGAGGGTTACGACATTAACCACGATCGCTATCCGGAGGGATATCAGGAGGCGCCACTCGCCTACGATGCAGTGTGGAGTGTGG
CTTTGGCTTTCAACAAGACCATGGAACGATTGACAACCGGGAAGAAATCTCTGAGGGATTTTACCTATACGGACAAGGAGATTGCCGATGAAATCTACGCTGCCATGAACTC
CACACAATTTCTGGGTGTATCGGGTGTGGTGGCATTCAGTTCTCAGGGCGATCGTATTGCTCTTACACAGATCGAACAGATGATAGACGGCAAGTACGAGAAGTTGGGTTAC
TACGATACTCAGTTGGATAACCTATCCTGGTTGAATACTGAACAGTGGATTGGTGGCAAGAAGACAAAGCATGTAACGACTGAACCGAATAAGAACGGAAATCCGTCCAAAG
ACGACGTTAATACAAGAGGATCATTAACAAGCTTAAGTGAGAAAAATGATGAATCCAAAAATCCATCATCATCTGCAAGTCTTCAGCAAGCTCCTAAGGTAGCACCAAAACC
GAAAAAGTTATCAATTTCGGATGCTGGTAAAGACACAGTAACCCAAAAAGTCAAAGAAAATGAAGAGCCAGGATTTTCTAAACAATTTGAAAAAGAATCTATTGGTGTGCGA
GCAAATAGAAACAGCAGTGCAACAAAAAAGGAAAATGAAAAGCTATTAGTAAAGACCGTGCCCGGAAAATCTCTTATAAAAGAATCTAACGATGAAAATGTGGAACCATCTA
GACGCACAAAAAGTCAGCCAGTTGGAAAAAAAGTTTACGAGGAATCTTCAACGAGGCGAGTTCGTGAGCCATTTGATTCTTTTGATAGGGAAAAGTACCTGTCTGATATGAT
CGATTATGATCGATCGTCCGAGGACGTGCCGGAAAAAAAACTATCTTGGGAACCAGACAGTACACTTCGACGTCGTTTTGTCAGCAATAATGAATATCATTCTAGTCTTGAA
GGTGAGGAGGAAGAGATCGATTTAGATTCAGTTGGATCGTCAACAATGCGAGGTAGCTATCGAATGCCACAGAATGCCGAGGATGAACGACCTGTTATCCTAGCTGAAATAA
TCAATATGGGAAAGTTGAAGGAGCTCAAGGAGCAACGCGAATCAACTAACGTATTTCGTAACAATCGTAGAGATGATAAAACCATTGATGAAGTGAAGGGTTCCAAGATCAC
AGGAACCAGCAAACAACATGTTGAAGAAGTTAAATATAATGAAACAGCAAGTAATGAAGATGATGACAATGCACAGGTAAAACATATAGGTACAATCACAAAAATCTTTAGC
ATTGCGCAAAGAATGAAAGAGAAGAAAAGCTCCAAAGAAAGTGAAGAAGACGAGAATAAGGACGAGAATAAAGATGTAAAAAAAGATGAAAATAAAGACGATGAACAGAAAC
TTCCCCCGGAACCAGAGATCGAATTGGAGGCCAAAAAAGCTTGGACTTTTCCGATCGCAGAGACATCACCTGATATCCCTGGAGAAAATCAGATCCTGCAAGAATACCAAAA
GGTTAAAATTGGCAAAAGAAACAGGAAATACAGATGTATAGGCATCAATACGGATATAAGCAGAAAATCTAAGATTAAAACATTTATCAGCAAGTCTGCAGGTGATGACTTA
GTAGTATACCATGATGGTGGAAGATTGCGAGATATAGGACAAACCACTGAAAATCTTAAGAATCAACGAAAGAATCATAAACACAAAACCGAACCAGATGATATACCCGTTG
ATATAGGACACAGCGATGATCGAGTAAGGGAAATTGGAGTGAATACTAAAAAGTTACCAAAAATCATAATACCACCAATTGCCGAGATGCATGTGCACAAAAATGGCAAATT
AAGGGACATTGGAACCAGCACAGATAAGCCATTTTGGCCAATCGATGACGGAACTGATGTAATCTACATGCATCCCATTAAGACTGATAGGAAGAAATTAAACAAACTGATT
GTGGATCCACCGCCGGATAATGGACCCTACAAGATGCCAACCAAAGAAGATAGACGAACGTACTACAAAGGTTGTGAATATCATTTTCCTGGCCGAACGGAATGGCGCAGGC
TTTTCTTCAACAAAATCCATGGAAAATACAAGCTACGCCGTCCAAGTCATTGGCTCTACACACTAGTTTTTTCCGTTTTATATATTTTGTTTGTCATTATCTTTTCCATGGC
CTGGTTCGATTTCATCAAAGATGACGCTTCCAGAAAGGTACCCATGATAAAGATGGCACAACCATTTATTAGTTTCACTCCCATCGGACCACGAACTAATCCCAAGGCGGTA
TCCTTTGATCCTAGAAACAGCACAGAAGTCATGGAAAAGTACGCTGGTATAATGGCTCTTCTGGAAAAATACGGCGACTATGGACACAATCCGCGTTTTGGAACGTGTACTG
CGAATGAAAAGTTTGGCTATCCAAGCGGCGAACCGTGTGTTTTTCTTAAAGTTAATCGGATTATTGGATTTAAGACCGAACCCTACATCAATTCCGATGAGCTTGTCAAAGC
CAAGATCGATGAGGTTGAATTTACGGCCTTAAAAAGGTTATTGGAAAATACAACAACAGAGGAGGGTCATTTGAATCGCACTTGGATAACATGCCGTTCAGATAAAGATAAG
AATGTTTTGATTGAATTTCATCCAGAGCCGGCGATAAGAACGGAATACACCGATATTGAAGAGAAAATCGAGTATATTGCGAACGAGGGTAAAAAGTCATTTTTCGGCCCAA
ACGATGTGAATCGAATCGTGGCTCTTAAGATCAAAAATCTCAAAGCAAATGAACGTGTCCATATAAATTGCAAAATAGTAATACAATCCTCGCATCCCGTTTGCAATACGAT
CATGTTATTTGGTGTCATCATCTGTCTAATATCTGTCATCTTACTGGGCATCGACGGACGCTTTGTCAGCCCCGAGGAATATCCAAAGATCTTTGATCCGCTGCAGCGTTAT
CTCGAAACATTCCCACTCGAAGATCCAGTATCTACTACTGATGATATTAAAATACGTCCAGAGCTTGAGCATTGTGAAAGTCAACGCAACTCCATGTGGTTGGGTCTTGTAT
ACGGCTTCAAGGGGCTAATCCTGGTGTTTGGCCTCTTTTTGGCGTACGAGACGCGCTCCATTAAAGTGAAACAGATCAACGATTCGCGTTATGTGGGCATGAGCATCTATAA
CGTGGTCGTCCTTTGCCTGATAACAGCTCCGGTGGGCATGGTCATTGCATCGCAACAGGACGCGTCCTTTGCCTTCGTTGCTCTAGCTGTGATATTCTGTTGTTTCCTAAGC
ATGCTGCTGATATTTGTGCCAAAGGTCATTGAGGTTATACGTCATCCCAAGGATAAGGCCGAATCGAAATACAATCCCGATTCAGCCATATCGAAAGAGGACGAAGAACGCT
ATCAGAAACTTGTTACCGAAAACGAGGAATTGCAACGATTAATAACACAGAAGGAGGAAAAGATTCGAGTCCTGCGACAGCGTCTGGTGGAGCGGGGCGACGCCAAGGGCAC
AGAACTGAATGGTGCAACAGGTGTCGCCTCCGCCGCCGTTGCAACAACTTCGCAGCCCGCTTCCCTCATCAACTCATCAGCACATGCCACGCCCGCAGCCACACTCGCAATC
ACACAAGGTGAGTAG
(SEQ ID NO: 176)

Start ATG: 26

MWNTTNTAGASRVSSPEYNAKKTEYLQKEIRVKWNYAPPPSNTPYDPAQAKGAAKKKKRLMKQLKHRPKEQRSRQRQRQKHRQRIKRRPRRRGSSSSKAGNNNNNNRNIAFL
NGNTQKKYCWQEFRTLPKGMRTDEGAGAWRDMTSDGAVTFWIFLLCLIASPHLQGGVAGRPDELHIGGIFPIAGKGGWQGGQACMPAARLALDDVNKQPNLLPGFKLILHSN
DSECEPGLGASVMYNLLYNKPQKLMLLAGCSTVCTTVAEAAKMWNLIVLCYGASSPALSDRKRFPTLFRTHPSATVHNPTRIKLMKKFGWSRVAILQQAEEVFISTVEDLEN
RCMEAGVEIVTRQSFLSDPTDAVRNLRRQDARIIVGLFYVVAARRVLCEMYKQQLYGRAHVWFFIGWYEDNWYEVNLKAEGITCTVEQMRIAAEGHLTTEALMWNQNNQTTI
SGMTAEEFRLVLLLYPLLKFLLQCAFFLRRHRLNQALIEEGYDINHDRYPEGYQEAPLAYDAVWSVALAFNKTMERLTTGKKSLRDFTYTDKEIADEIYAAMNSTQFLGVSG
VVAFSSQGDRIALTQIEQMIDGKYEKLGYYDTQLDNLSWLNTEQWIGGKKTKHVTTEPNKNGNPSKDDVNTRGSLTSLSEKNDESKNPSSSASLQQAPKVAPKPKKLSISDA
GKDTVTQKVKENEEPGFSKQFEKESIGVRANRNSSATKKENEKLLVKTVPGKSLIKESNDENVEPSRRTKSQPVGKKVYEESSTRRVREPFDSFDREKYLSDMIDYDRSSED
VPEKKLSWEPDSTLRRRFVSNNEYHSSLEGEEEEIDLDSVGSSTMRGSYRMPQNAEDERPVILAEIINMGKLKELKEQRESTNVFRNNRRDDKTIDEVKGSKITGTSKQHVE
EVKYNETASNEDDDNAQVKHIGTITKIFSIAQRMKEKKSSKESEEDENKDENKDVKKDENKDDEQKLPPEPEIELEAKKAWTFPIAETSPDIPGENQILQEYQKVKIGKRNR
KYRCIGINTDISRKSKIKTFISKSAGDDLVVYHDGGRLRDIGQTTENLKNQRKNHKHKTEPDDIPVDIGHSDDRVREIGVNTKKLPKIIIPPIAEMHVHKNGKLRDIGTSTD
KPFWPIDDGTDVIYMHPIKTDRKKLNKLIVDPPPDNGPYKMPTKEDRRTYYKGCEYHFPGRTEWRRLFFNKIHGKYKLRRPSHWLYTLVFSVLYILFVIIFSMAWFDFIKDD
ASRKVPMIKMAQPFISFTPIGPRTNPKAVSFDPRNSTEVMEKYAGIMALLEKYGDYGHNPRFGTCTANEKFGYPSGEPCVFLKVNRIIGFKTEPYINSDELVKAKIDEVEFT
ALKRLLENTTTEEGHLNRTWITCRSDKDKNVLIEFHPEPAIRTEYTDIEEKIEYIANEGKKSFFGPNDVNRIVALKIKNLKANERVHINCKIVIQSSHPVCNTIMLFGVIIC
LISVILLGIDGRFVSPEEYPKIFDPLQRYLETFPLEDPVSTTDDIKIRPELEHCESQRNSMWLGLVYGFKGLILVFGLFLAYETRSIKVKQINDSRYVGMSIYNVVVLCLIT
APVGMVIASQQDASFAFVALAVIFCCFLSMLLIFVPKVIEVIRHPKDKAESKYNPDSAISKEDEERYQKLVTENEELQRLITQKEEKIRVLRQRLVERGDAKGTELNGATGV
ASAAVATTSQPASLINSSAHATPAATLAITQGE*
(SEQ ID NO: 177)

Name: GABA B receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384569
TGTGGGTGGTTCGCACGCCCGCCATTTCCGGCGCGGCCTCGGATAACGTCCTCGATTACGCCCTGCCGCCCAACACGCCCACTCCGAAGGACTACGCGCACAAGTTCCTGCC
CGAGGGCAAGGATGTGGTACCCGGAAGCTCTTACATACCGCTGCGGTTGCTGGTGATAAAGCGCTGAAGCTGATATATGAATCTTGGGCGGCGGCTTGCCTTTTGTTGTTCG
AAAACAAGTTCCGGCGGAATGCCGATACCCGACACTTAATGCCCAACCAGTAGTTGGCTAATTGGCTTTAAGCTTACTGGTTCCTAGCGCTAGTCTAGCAATAGTTTAGTTG
TTTATGTTAAGGCCCAAGTTCGCTAATCTAAATTAAGCCACCACATCCCGGGGTCCGTGACTAAGACTTCCGACCTTATCGCATCCAATAAATCCAAAACTAAGACTCAGCC
ACATGCCGTCGTAATCTATGAAATACATTCACCAAAAACTAAGTTGTCATCTCATTTACATAAATTAGAAGTCATAAATTTGAAGTGAGCGGGGAACTACTTGGCTCGAAAC
GGTTAGAAGCTGAATTCAAAACCGGTTAACGTTTATCTGTCAGCTGTAATCAAAAGCTTAAAGCTTCGCTTGTTTGGTTTAGTCTTTGACTAAGTTTGGAAAGCAGGACTAA
AAGCTTTTATTTGCCGAAACTAAAGCTCGGAGGTATTTATTTATTCACAAGTAAAATGACAAAACTTGGAAGTTCTTTGGTAATGTAATGAACGCCAATATTTACATTCGGG
TTTTATTGTATTTGAAAAGTCTGATGAATCATTATCTTGATGACTGGCCCCTAATTTACTTTTCATGTTCTACGCAAGGGGTATGATAACTGGTTGACATGGATACCAAGAT
TAGAGCTTGCAGACTTTATAATCATCCAACAAACATTATGGTGCATTGTATACCATTTATGTTATGAATGAAAGTGGTTCAGATAAGTTACATTTTATTTTATCTCACACAT
CTAAGTTTTTGTATTGCTGCACAGGAAAATAAATCAAGATTATCCACAATAGTTTAAAAGTCAATTTATAGGTACTATGTAGAATTTATTTACATTTTGTTGTACTAATGGT
AACTTAACTTAGGTTTTTAAATACTAGTACTGTTTGTTGATGAATGGCTTAGGCCTTGCAAAATCGAGTCAGTGGGCATTTGTATAGTTTGCTGATTTCTCTTCTTGCTCGA

ATTGCAACAGAGACCCAACCAGGCCCGCTGGTTCTCTTTGTTAAATACTATAAAGTAAAGGAATAAAACAAATCCCTGTAAAGTGGCTGTCAGGCAGAATAAATAGGCGAAG
ATCCTTCCGAAATCGAAGTATGTACACAATCCAAAAATCCATGTTATTCCCAAAAGAAAGAAGAGCAACACAAACAGGCCGAGTTGCTTGAAGATAAGGTCACGTTTGAACA
GGGCTGTATAGACACTCCAGGAGATGTATCCCACCAATATTGCATTGGCCACCGTGATAAGGCCAATAGGGAGTATGACGCCCAGACTTAAACCATATCCAGAAGGATAGCA
GAGGATGGGATAGTCCATTGAGGAATTATTGGGTCTGTAGGAACCGGGCTCCAGAAAAACCACCAACAAGGTGGGTATAAGTGGCAATGTCCAGGCCGCCACAGCCGACATT
AGTATATAATGTCGTGGGTGATTAACTCCAATGACTCTCACATATCTTTGATACTGCAGAAATCCAATGATAAACATCCAGCTGAAGACAACCAAGAGTAGATATTGCATCA
CGGCTCCAACCAGAGTACATCGTTCCGACTCAGATTGTTCCAGCTGCTCCAGCAAATGGCTTTGACTCAGGATCAGGAAGAAGAGCAACTGAAGACCCAGGGCAGCACACAG
ATTCAGCAATATCTTCGTGGAGGCTAATGTCCGGAAGCTCTTAAACACGGCTGCAGTTATAAAAAATCATTAGGAGGCCCAACAGAGAAAGGGTCAAGCCGACATTGGTGATT
ACATCTAAGGAGTGATCATCTTCATTGGTGTCCAGACCTGCCTGCATTTTGCTCACTCCCAGTAGAAATGTAAACTGGGTCAGATGATCCGCATGGCAGAGGATCACTGGGT
GCGATGAGGCTTCCAAGCTGCCACCACCGCTGGTGGATATTCCATCGCTAAGCCAGGTCTCATAGTTCCAGTAACCACATCCTCCCTCAAATTGAGTTTCATTCGCTTTTGC
ATTACGCAAGAAAAATGGCAACTTTCCAGGTAGCTTGGTATCTTACGAAAAGATTTTAGATTAAAACATATATGCTATAATTTTCCCTCAACCCACTCACCATTCAATCCTG
GTATTGTTATGGATATAACATTACTAGTTGGCCGTCGTTTGATTTCCTCTTCCGTCTCCACAAAGAGTGAATCTCGAGTATAGACCTTAAAAACTAAGTAAGAGGCTCCCTT
TCTCTTCAGTTGGGTCCATAGTTTTTCGGGAATAAAGACCGCACTTTCTAGATTCTTCATGGCCCTTATATCTTCCTCATTATCCGATAGGTGAAGCATTTCGAAATGCGAC
ATATTGCCACCAGAGAAGTTCACTATCCCACTCACATTTCTCACTTCTGGATCTACAAAGAATACTCTGAGTTCGCTCGTCTTTAGAACAATGACACCCATATCAAACAAGG
GATGGAATGTGAACTCAGTGCTGTTTGCAAAAATAGTGGTGTTTAACTTTAAAATTGTTGGACTAAGCTCATCCAAATAATCCTCGAAGTTGCTAAGGAGTGCATTGGTTGC
ATTAGCGTCTGCTGAAACACGAAGGGTTTCATTATCTGTAGACATAACTTTTTGGCATATCTTCATCAAGTCCCCTACAACTGCAGCATTTCTTTCCGGTTTTTGACCAGC
GAATCGAATATCTCCGAAATACTGAAGACATCGGCGGCGGCCAGTTTCTCCTGAGTTTGAGTCAGGAGGTTACTCAGTTGGCTAATCATGTCCGGAGGGCTCTTCACAGCGA
GAAGGTTTATACTCGAGCTGAGTTCCGGTGCCGACCTACAGGTCACGTGATCCAAGCTCTGCCAAACTCTATTCCCATTGCGAATTTGGCACCTCCGTCGGATGGGAATTCC
CCGTTCGTTCAGGCATAGATCCCTTAGATTCGCCTTTTGCCCAACTCTCGTTTTCTCCACGCGATTGATTCGCTTTTCCTTGCCTTCATCCGTTTCGTATTCCAGTTCAAAG
TCCTCGGCCTTGCAGAATCCTCGGTTGCAGTCCTTGTAGGGATTGTGCGACAGCTGACCATCTGCATCCACGCAGAGGAGGAAGCGCAAGTTCTTCAGGTGGCAGTAGTCCA
GACACCGGGCCACATTCGGCTGATCCAGGGCGTACAGGGGGGATTTGTACCTCCAGATACGAGTATTATATAGCCAGCAATCCACCAGAGAAACAGCAGCATACTGCTATC
TATTTTTAACTAATTTAAGCGCAGGCTGCGATGTCGCACTCCCGAAGCTCCGCGATCGACTGACGAGGCGGAGATAAAATGCGAGTGGCGATAAAACGAGTTTTCCTCTCCT
CATCCCTGTAAATAGGGCCTGCAAGAATTGCCACTTTCCTCACTGTTCTAGCTGGGTATTTGAATTGCTGGTCTACTTCGAGCAGAGCTCAAGGTCAACTGGTTATCTGGCC
GACGATAAGAATTGCAGCATCTTCAAAGGTCGAACGGTGCGTTCTTTTCAAAATATCTTTGAATATAATAAGTGAAAGTCTCGAAAATAACACGAACCTCTTTAATAATGGG
TATTATTTATACACAGTAGTTCTGTTTACTTTGATCACATTTTTTATCAAGGGAACTTCCTAATTAAAAACAAAAAAAAAAAAACGAATAGCCTGATAAAGAGCTAAATGGT
AATTATCAAAATTATATATACCATACACCCAACGATAATGATAATTTGTTTATCGCAGTAAAGTGAGATTCTCTGAATTGATAAAGGTTAATTGCCGTTTTTCAACACCATT
GCATGTATCGATAAAACTAATAGGAATTACAAATAAATTAAAAATCTGAATTTTCTTTTTTGCTCTGGGTAACAATGGCTGCCATTTCTGAATGTTTCTGTTATTCTCTCTG
ATTTTAAAAGTGCAACGATACTTAACACGAACTTTGAGTTGCTAGTCGCCATAAGCATATGAATGCATCAATCGGTAGCTAGGGGCTGTCAACAACTACAATATGTATGATA
AGCAGCCACCTTTAGATACTTTCCTTACTAACACCTAGAAGGTAGAGATAGAGGTATTCAAATATTTGCATTCAGATATAAAATAATGTTCTCTTGTTCAAGCTGTGAAATC
TTTATCTGTAGAGATAAAGTGAGCATTTCAATTGAAGAACCAATTTGTATATTTTTTAGATTTCACCGAACGCATGAATAGAATTTGTCTGCTA
(SEQ ID NO: 178)

Exon: 3463..2341
Exon: 2281..1491
Exon: 1024..1001
Start ATG: 3463 (Reverse strand: CAT)

Transcript No. : CT35672
ATGCTGCTGTTTCTCTGGTGGATTGCTGGCTATATAATACTCGTATCTGGAGGTACAAATCCCCCCCTGTACGCCCTGGATCAGCCGAATGTGGCCCGGTGTCTGGACTACT
GCCACCTGAAGAACTTGCGCTTCCTCCTCTGCGTGGATGCAGATGGTCAGCTGTCGCACAATCCCTACAAGGACTGCAACCGAGGATTCTGCAAGGCCGAGGACTTTGAACT
GGAATACGAAACGGATGAAGGCAAGGAAAAGCGAATCAATCGCGTGGAGAAAACGAGAGTTGGGCAAAAGGCGAATCTAAGGGATCTATGCCTGAACGAACGGGGAATTCCC
ATCCGACGGAGGTGCCAAATTCGCAATGGGAATAGAGTTTGGCAGAGCTTGGATCACGTGACCTGTAGGTCGGCACCGGAACTCAGCTCGAGTATAAACCTTCTCGCTGTGA
AGAGCCCTCCGGACATGATTAGCCAACTGAGTAACCTCCTGACTCAAACTCAGGAGAAACTGGCCGCCGCCGATGTCTTCAGTATTTCGGAGATATTCGATTCGCTGGTCAA
AAAACCGGAAAGAAATGCTGCAGTTGTAGGGGACTTGATGAAGATATGCCAAAAAGTTATGTCTACAGATAATGAAACCCTTCGTGTTTCAGCAGACGCTAATGCAACCAAT
GCACTCCTTAGCAACTTCGAGGATTATTTGGATGAGCTTAGTCCAACAATTTTAAAGTTAAACACCACTATTTTTGCAAACAGCACTGAGTTCACATTCCATCCCTTGTTTG
ATATGGGTGTCATTGTTCTAAAGACGAGCGAACTCAGAGTATTCTTTGTAGATCCAGAAGTGAGAAATGTGAGTGGGATAGTGAACTTCTCTGGTGGCAATATGTCGCATTT
CGAAATGCTTCACCTATCGGATAATGAGGAAGATATAAGGGCCATGAAGAATCTAGAAAGTGCGGTCTTTATTCCCGAAAAACTATGGACCCAACTGAAGAGAAAGGGAGCC
TCTTACTTAGTTTTTAAGGTCTATACTCGAGATTCACTCTTTGTGGAGACGGAAGAGGAAATCAAACGACGGCCAACTAGTAATGTTATATCCATAACAATACCAGGATTGA
ATGATACCAAGCTACCTGGAAAGTTGCCATTTTTCTTGCGTAATGCAAAAGCGAATGAAACTCAATTTGAGGGAGGATGTGGTTACTGGAACTATGAGACCTGGCTTAGCGA
TGGAATATCCACCAGCGGTGGTGGCAGCTTGGAAGCCTCATCGCACCCAGTGATCCTCTGCCATGCGGATCATCTGACCCAGTTTACATTTCTACTGGGAGTGAGCAAAATG
CAGGCAGGTCTGGACACCAATGAAGATGATCACTCCTTAGATGTAATCACCAATGTCGGCTTGACCCTTTCTCTGTTGGGCCTCCTAATGATTTTTATAACTGCAGCCGTGT
TTAAGAGCTTCCGGACATTAGCCTCCACGAAGATATTGCTGAATCTGTGTGCTGCCCTGGGTCTTCAGTTGCTCTTCTTCCTGATCCTGAGTCAAAGCCATTTGCTGGAGCA
GCTGGAACAATCTGAGTCGGAACGATGTACTCTGGTTGGAGCCGTGATGCAATATCTACTCTTGGTTGTCTTCAGCTGGATGTTTATCATTGGATTTCTGCAGTATCAAAGA
TATGTGAGAGTCATTGGAGTTAATCACCCACGACATTATATACTAATGTCGGCTGTGGCGGCCTGGACATTGCCACTTATACCCACCTTGTTGGTGGTTTTTCTGGAGCCCG
GTTCCTACAGACCCAATAATTCCTCAATGGACTATCCCATCCTCTGCTATCCTTCTGGATATGGTTTAAGTCTGGGCGTCATACTCCCTATTGGCCTTATCACGGTGGCCAA
TGCAATATTGCAATACAAAAACTTAGATGTGTGA
(SEQ ID NO: 179)

Start ATG: 1 (Reverse strand: CAT)

MLLFLWWIAGYIILVSGGTNPPLYALDQPNVARCLDYCHLKNLRFLLCVDADGQLSHNPYKDCNRGFCKAEDFELEYETDEGKEKRINRVEKTRVGQKANLRDLCLNERGIP
IRRRCQIRNGNRVWQSLDHVTCRSAPELSSSINLLAVKSPPDMISQLSNLLTQTQEKLAAADVFSISEIFDSLVKKPERNAAVVGDLMKICQKVMSTDNETLRVSADANATN
ALLSNFEDYLDELSPTILKLNTTIFANSTEFTFHPLFDMGVIVLKTSELRVFFVDPEVRNVSGIVNFSGGNMSHFEMLHLSDNEEDIRAMKNLESAVFIPEKLWTQLKRKGA
SYLVFKVYTRDSLFVETEEEIKRRPTSNVISITIPGLNDTKLPGKLPFFLRNAKANETQFEGGCGYWNYETWLSDGISTSGGGSLEASSHPVILCHADHLTQFTFLLGVSKM
QAGLDTNEDDHSLDVITNVGLTLSLLGLLMIFITAAVFKSFRTLASTKILLNLCAALGLQLLFFLILSQSHLLEQLEQSESERCTLVGAVMQYLLLVVFSWMFIIGFLQYQR
YVRVIGVNHPRHYILMSAVAAWTLPLIPTLLVVFLEPGSYRPNNSSMDYPILCYPSGYGLSLGVILPIGLITVANAILQYKNLDV*
(SEQ ID NO: 180)

Name: HE6 Receptor-like 2
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384028
CAGCTAGTAGTGAATAATACTTGAGTTTTTGAGTTTGTTGGGCTGGGTTGAAAAAAAAAAAATATGCACTTTAGCCTCTATTAAGCATTTTCCTATGGTTTTCACTCTCGTTC
GGGAACTGTGTGTTTTGAAAATGTAACAATTTGCGCTTTCGTCCTCGTACATTGTGTGTATTTACAGAAAGTCTAGCCCATTTTTAGATTTTGTATGGGAATGAGTGGAAAT
TACAAGTGTATAATATATATAATACAAATACACTTTAGTTGGTAAGATTAGCAAGACGTAAGACGTACGATAGTTAATGAGGGGGGAATTGCTAGTTTTATGTTAATGCCTATA
AGCTCAAATTTACAACTGATTATTCGCTGATCGAAAATGATCTCAAGATCTCTATCATTTGTTTAGCTTTAGTGTGGATAGTGTTGAAAGGTTGACTTGAAAGCAAATGGTT

AGTAACTAACTGTCGAGGAACTTAGAGGTAAACGAAAACGCGTGTTTTAGGAATGGGTTTGTTTTTCTTAGTGTATGTGGCCTTGTTTTTACAGTTATATCGTTACCTACAC
AGGTGGGCTCTGTTTTACCGATCTAAACGTTTGAACAACTTTGATCTATCAGTGATCCTAATCAATCGAAATCAATCGAACTCAAGACTGACAGAAATCAACGTAAACTAAA
GTAAAATCGTCTAAAAATTTATCAAAACTTTATAAAACTTCTTGGCTGCTTTCAAGGTTCCTCACTTTACGTTTCATTGAAAAATAAAGTAATATTTGTTATTGTTTGTGTT
TGGTGAGTGTTTGTTGCATCTCCATTTGCATCAATCATTTTCCTGCTGGTTTTTTTTTCTTTCTTTGCGAAATTAAAGTAATTCTTTTATTGTTTGTGTTGTTGCGCTATA
TAATAACTATTTCTCACATATTATCTACACATACCAACGTGTATCGTACCGTTTAATACCTATATACAATTCTCAAATATCCGCCATGAACCAGCAAGGACTCCCTACTTCG
AATGCGAGACCACGGAGTGGGAGACCATCGAGTGAGAGAGCATTGAGTGGGACACGCCGGAGATTTCCGCATAGGTATCCGTGGGTCCTTCGGCACTGCACACCGAATCCTC
GGGCGTCCTGGAAGGTTCCGCCTCGGTGATCTCGATGTCCTGATGATGCTTGTTGCTGCGGTGTTTGGCACTTAGAGCCTGCGAATTTGAGATATTTTAGTGGGTCACATGG
ACAACTGGTGGATATGTGTTTTCTCGTTGTAGTTTCGTGTGCAATCTGTTTTCGATAAGACTTATGTGGTGCTATCTATGGGACGGACACATGCACACACACACACACACAC
GCACAAGCTGGTGGTTTCGGGTTCATTGGGGAACTTTGGGTCTAAGTTTATAGACAAACGTGCGAGAAATTCTTTTGATTTAGTCGTTTTGTTTTGGATTTTAGTTTACATT
CACAACACAGATATTTTAGCAGGTTAGGTTTAGAGGTTTATGCTAAAAGCAACCGAAAATTCTTACCTTATCCTTGCTGCCCTTTTTCTACATTGAACATTGGTGGTTATTG
TTTTGTTTGGTTTGGTTGTGGTGGGTTTCGAATATATTGTACATAGTCATGACAAGAACAATGATATTTATAGATGTATTAATATAAGTGGCAAACAGTACTGCGAGTTCAC
ATTGGAGGATCAAGAGAAGTACAGATTCATAGACGGTTAGAGAGACGGGGGAGAGAGAAAGATCTAGGGAGCGTGGGGATGGGCAGCAAAAGGTTTGGTTTCGGATTGAAGA
ATAACAGCAATATCGATCTGAGGCCTACAGCCCGGTTCCCCATTCTGGCCACGTCAATGTCATGTACAGTGAGCACTATGGCTGGTGGTGGACTACGCTGCCGTTGTGGTGA
CCCCGACTCACCTGCAGGGAGAAGCGACGGTGACCCGCCTTGCGTCCGCCACGTCGCTTACTGATGTGCGGATTGTCCTGCCGCAGAGTCTTGTTCTTCATTATCTCCAATT
GGGTATACAGTCTGGAATCGAAAATGTATAGAAAATAGAGATAATAGAACATATAGATATGCTTCGTACCTTTTGAGTTCGGCGCGTATGTCCTCGGGACTCATTTCGGATA
TGGTCAGCTCACCAATGTCCGGATCGCCCAGACAGAGCCCGGCATATCCTCCGCCCAGCCGCTGGCCGGCGTCGTGTGACGTACCCTTGAAAGCATCCACAGGTAAACGTAT
TGATAGATCATACGCTAGCGAACGAACCTGCAAGCTATGAAGATGCAACAAACACCGCAAACCACACTCAGGAAGTCAAACGGAAAATGGAGAACGGAACGGAAACGGAAAT
GGAAACGGAACTCAAAGGTAAGAGAGAGCAAAATCGCATGGTATAGATCTGTGCTGAAGAGTTATGTATAGTTATATGTATATAGTATGTGTATGGGTGGTCTGTGCTTATT
GGTCGTGATTGTATCTTTGGGGGTGCTGGAAATGAACGAAAAGAAACAAATCAAATTAAAAAAGTAACTCAAAAGGTAAACAAGATGGACGATAGCTATAAGAAGACTACTT
AGACAGAGAACTTCAGTTTAAGATTATTCAGTTGAGGCACAGAAAGACAGAGAAAGTTACAGTGAGGCAGAGTCTTGGATACACTAGTTTGGAATTTTTTTTTCTGGTATGC
ATATTGGAAATTTTGAAAGGAAACCGATTTTTGTCATTTGTTTGTTAGTAGTTAGATGTTGTGTTTTCTTGTTAAATTCATGCAAAAATCAACACTAAACATGCAAATTAAA
TGGCGTGCTTATTGAGTTCAAGTTAAAGTGCAATACATTCAGGCATAAATAATAAGAGCTAAAAACAGTCGAACATTCTCTGTTTAAACGAGACACACTTGCAGCAACTTCA
TAGCTTGACTAAAGTTTCAATTCAATGTTAGCTTATTACTAAAATTATGTTAATTTTCTCGGTGAAAACCAGAACAAGAACTCAAAATATAACCATAAACAAATGAAATATT
CAAATTGTATTTTAATTAGCAAAAGCAACCAAGCTTTTGAGTGGATGTCCACATTTTAATACTTCTTGTGTTTTGGCTAGTTATTTAACAAAAATCGTTTCCTCAACCAATG
TCAATACCAAGTACGCATTTCGATGACTATAATGACGACCACGCCCACGACAGACGAAAGGCACAAAATGAACGGGTGTGAATACTCTAAGCATATGAATTATATTCGTTTG
GTGCCCGACTAGGACAAAACACTTCCTAATTACATCATATATCTCAATGTGTAGAACACAAAAGCGTACTCAAACCAAAAAAAAAAAAAAAAAAAACCAAAGCAAAGTAAGTA
AATCATATCGTTTCGAACTCAAATCGGTTACCTGCTTGTGCTGGTACCACAACTTTGGCACAAATATCAGACCCAAGGCGAAACTATTCGTCAGCTGGGAGCGGATGAACAG
GGCCAGCAAGATGGCGCTGGGACTCATTTCCGGCAGGTAGACAAAGCGCAGAAAGTAGAAGCTGGACGAGACCAGGAACTCCAGCGTCAGTGCGGTGACCAGAAATTGTCGT
TCCTTCAATGGGACAAAAGGAAATTACCGATTAATTCATTTGACTTGACTTAAATTCCCCTTCTACTTAGTTAATCATCTAAGCCAATCTCTTGCGCTTAATTGATAAAGAA
AGATCTGCAATCGATTTGACTGTACATCTTACCCGGAACTGGGTGTTGGCATTCCGACTGGCGATGGACAGGTGCAGTCCAAAGCACAGGATGAGCATCTCGCTGGTCTGCG
TGACCAGTTCCCACTTGAGCGGATGGCAGGTGTTCGTGTCCGCCTCCCTGAGGCTCTCCAGCTGCGCACTTTCGAGAAGATCCAACGACGAGGCCGTAAAGGCGGCCATGTA
GCAGATGACGGCAAAGACCATGGTGCCCAGATACTTGAGCAGGTCCACGTCGCGCAGCACCCAGCGATGCGCTTTCCGGGTGCGAAAGTCCACCAGGTGGCGGTACAACTTC
AGTATGATGGCGCCGTAGCAGGTGATGAAGCCCAGCTCCCGAAGCCAGGGCTCCAGAAGGCAGCGCTCGGTGGATGCGGGAAAGAAATGGACGGCAACCTAAAGATGAATAT
GTAAATACACACAGAAGTTGAAGGTGACCATCGCCAGGAGCACTCACAGATGCATAAAGTAAAACAATGCCCAGCAGTATCGTCTCCAGCACAGTCCACATACCAGACGCAA
TGGCCTGAAAATGATGAAAGGATTACTCAAAGTTTCCCAACTTTAGGATATGTAAACTAAATACCTTGCACTTTCGCTGCCGGAAGACAATCACCCCGAGGACAATGCAGCA
CAGGATGCAGGCGCCCAGTACGATGGCCACCAGGAGGCGCAGACAGGCGTCCACATTGAGCACCTCCTCCTCCTGAAAGGTCAGACAGACGCCGTTGCTATCGCAGTTAGTG
CATCCGCCGGGACACGGAATGCACGAGTAGTTGTCGTAGCCCTCGGACAGTTCCACTCGATCGCCACGGAATCCCTGGAGCGTGGAATTGGGCAGGTAGTAGGATTCCCGGC
ACAGACAGGTGTACACATCCCGGGTGGCGGCGGGTTTGTTCTCGGTGAGCAGGCAGAAGGTCGTATTCCGATCGCAACTACAGAGATATGGTATTATATATATTATCCAGAA
CTAATGTCACGAATTATCGTATAAAGTTTGGCCACAACTCACCCATGACGGCGACCGAAAACTTCCTCCAGGCCATCGTTGCACACATCCTCGTCGGCGGCAATGAAGGCAG
CCGCCACAACTCTGAAAGCATTTAGCATAAATTACAAAGTTGTGCCAACCAGCTGCACATCGATGCGACTCACTTGATCCTGTGCTCGCTGAAGGATATGCGAAATGGCCAG
AGCCAGCGACGCGTCAACAGGTTGCAGTCGCGAAATGGAGCGCCCCACTCCTGGGTGCCGTCCGTGCTGTTCCAGGCCGCACTGTAGTCGGGTAAAATAATACATTTATGGG
CAACTTTCCAAGATTATACGGCTTATGGTCTATTCAAATTAAGCGATGCACATTATTTCGGCAAGCTACGGTGCATATTGAATTTAATTGTATTTATCTTTTAGAGTAATTA
GTATTTGCATTTCAAAGCCATAAGATAATTGATTTTAATTTGTTGCTAATTAATGTCGCACACTTTTTTGGGACACAACCAGTCTTCGAAACTTACCCCAACTCCCGCCAAT
AGGAACCCAAATATGCCTTATTGTGCACCTGATCCGGTGGTATCTGCAATAAGTTAAATCGATTAGCCCATGGCCCCATTATCGCTGATTAGCGAGTGCATACCGTAACGTT
ATCCCTGAATTTAACGACCACCGTGTGAATGCTGGGCGCGGCAATGGCGAATGCCAGCAGATTGTCGTCCTCGTTGAGGGCGTAGCTGGCGATTCCATTGGCCCGGCGACTC
ACCTCCGCACGTGTCGGCTCATCGATGGCAATCGGTGTGTTGGCATTTGGGAATTTGCCCGATAAAAATTCAATAAACTGGCGTGCGGCATCCTCGGCGATGTCGGTGTAGT
CGTAAAACTGTTGAATTCATAAAACCAGATCACAATTAATTGCGTTGTGTCATAAATAATCATTTCAATGTATATGAAATGGGCATATAAATATCATTTAAATCGAATTCGA
TTTAAATACACAAAGAAAAAGTTCAATAAAATTAACACAGCATAAATTTACGCTTTGCTAAACTCAATTTATTTATACCACACTATATAGGAAACGGAAAACGTGTGCTTTC
TAATTGAAGATATTTTCGGAAATTGCGAAAAAATTACGTATATGCCACGTATATTGGTCGCCACGTGATTTAAGTGACCATGCGGCACGGCCACTTCCCCCGACTTGGGACC
ATTAATTATTCACCTGGCGCACAGGACACGTTTCACAGCAACACCTTGCGGCAAGAGTGCGAAAAGTTTTGCTTCGTTTCGACTGTGAAATAATAGGATTTAATACCCTCGC
CGTGCTGTGGGACTGCGGGAAGGGTTTTTAATCCAAGTTCGTTCAGCTCCTAACCGACCAAATGTTTATACATAATTTACACGTTACAAGAAACTCTCGAGATCTACGCACC
AAAGTGTAGGAAATGGTAGGGCCACTTTGATGACATCATTATGGCTTTGGGAAACAAAAGTAACCACAGAAGGGAGCAGGGAGGGGGTTGCTGGTAGCGGGGAAAAAGGACA
CGTTGTTTGGGCGTGAGTCATGTAGTTGGGCTGGCGGCACCTTTTGGCTTTTAAAACAAATATTTCCCATTAGCTGGGCTTTCGGTGATGAGCGACCACATTGCGTATACGC
CACCGAAGCCACCAAATAAATTGCCAACAAGTAACATCGATATACAAACGTTTGCTGGATGCTCCTCTACCATGTGTGTGTGAGTGTGTTTCTTCATTTTTTTTCCTGCCAA
ACGTGAGCCAAAATTTACACATCGAGGCAAAAGCAATCAGTATAAGGCAGCTGGAGAAATGGCACGAGACCTGCTCCCAGATAAAGTCATTATGTAGTGCTCCTAGCACTCG
CGCGCAGCCATAATCATAAACATTTGGTATACTCTTGCATGCTAGAGAAATATTGTAGTTGGGTAAGTGTTTTCAGATATTAAACTAAGTAAGGAAATATTTCTATTTTTAC
ACCATTATTCATTGTTTTTTCTTCAGGAGTTATCTGAACTTTAAGTAATCCCACCACAGTTCTTGGAAGCAGTAGAATGTCTTCATTACCGCCAAATCCTTTTGATATCCTT
TTTCCCAAATGTCCGCAATTACACACATACCCTCGCACTAGCGCCCCTCGTTTTTTCAGGGTATAAATGTATGGGGAAAACACTTGTAAATAGCGGCAGCAGCCAC
(SEQ ID NO: 181)

Exon: 5938..5903
Exon: 5607..5368
Exon: 5307..5249
Exon: 5009..4890
Exon: 4827..4747
Exon: 4669..4321
Exon: 4260..4192
Exon: 4130..3729
Exon: 3586..3392
Exon: 2213..2086
Exon: 2027..1916
Exon: 1543..1523
Exon: 1198..1001
Start ATG: 5938 (Reverse strand: CAT)

Transcript No. : CT35779

82/89

ATGGTCCCAAGTCGGGGGAAGTGGCCGTGCCGCATGTTTTACGACTACACCGACATCGCCGAGGATGCCGCACGCCAGTTTATTGAATTTTTATCGGGCAAATTCCCAAATG
CCAACACACCGATTGCCATCGATGAGCCGACACGTGCGGAGGTGAGTCGCCGGGCCAATGGAATCGCCAGCTACGCCCTCAACGAGGACGACAATCTGCTGGCATTCGCCAT
TGCCGCGCCCAGCATTCACACGGTGGTCGTTAAATTCAGGGATAACGTTACGATACCACCGGATCAGGTGCACAATAAGGCATATTTGGGTTCCTATTGGCGGGAGTTGGGT
GCGGCCTGGAACAGCACGGACGGCACCCAGGAGTGGGGCGCTCCATTTCGCGACTGCAACCTGTTGACGCGTCGCTGGCTCTGGCCATTTCGCATATCCTTCAGCGAGCACA
GGATCAAAGTTGTGGCGGCTGCCTTCATTGCCGCCGACGAGGATGTGTGCAACGATGGCCTGGAGGAAGTTTTCGGTCGCCGTCATGGTTGCGATCGGAATACGACCTTCTG
CCTGCTCACCGAGAACAAACCCGCCGCCACCCGGGATGTGTACACCTGTCTGTGCCGGGAATCCTACTACCTGCCCAATTCCACGCTCCAGGGATTCCGTGGCGATCGAGTG
GAACTGTCCGAGGGCTACGACAACTACTCGTGCATTCCGTGTCCCGGCGGATGCACTAACTGCGATAGCAACGGCGTCTGTCTGACCTTTCAGGAGGAGGAGGTGCTCAATG
TGGACGCCTGTCTGCGCCTCCTGGTGGCCATCGTACTGGGCGCCTGCATCCTGTGCTGCATTGTCCTCGGGGTGATTGTCTTCCGGCAGCGAAAGTGCAAGGCCATTGCGTC
TGGTATGTGGACTGTGCTGGAGACGATACTGCTGGGCATTGTTTTACTTTATGCATCTGTTGCCGTCCATTTCTTTCCCGCATCCACCGAGCGCTGCCTTCTGGAGCCCTGG
CTTCGGGAGCTGGGCTTCATCACCTGCTACGGCGCCATCATACTGAAGTTGTACCGCCACCTGGTGGACTTTCGCACCCGGAAAGCGCATCGCTGGGTGCTGCGCGACGTGG
ACCTGCTCAAGTATCTGGGCACCATGGTCTTTGCCGTCATCTGCTACATGGCCGCCTTTACGGCCTCGTCGTTGGATCTTCTCGAAAGTGCGCAGCTGGAGAGCCTCAGGGA
GGCGGACACGAACACCTGCCATCCGCTCAAGTGGGAACTGGTCACGCAGACCAGCGAGATGCTCATCCTGTGCTTTGGACTGCACCTGTCCATCGCCAGTCGGAATGCCAAC
ACCCAGTTCCGGGAACGACAATTTCTGGTCACCGCACTGACGCTGGAGTTCCTGGTCTCGTCCAGCTTCTACTTTCTGCGCTTTGTCTACCTGCCGGAAATGAGTCCCAGCG
CCATCTTGCTGGCCCTGTTCATCCGCTCCCAGCTGACGAATAGTTTCGCCTTGGGTCTGATATTTGTGCCAAAGTTGTGGTACCAGCACAAGCAGGGTACGTCACACGACGC
CGGCCAGCGGCTGGGCGGAGGATATGCCGGGCTCTGTCTGGGCGATCCGGACATTGGTGAGCTGACCATATCCGAAATGAGTCCCGAGGACATACGCGCCGAACTCAAAAGA
CTGTATACCCAATTGGAGATAATGAAGAACAAGACTCTGCGGCAGGACAATCCGCACATCAGTAAGCGACGTGGCGGACGCAAGGCGGGTCACCGTCGCTTCTCCCTGCAGA
AAAAGGGCAGCAAGGATAAGGCTCTAAGTGCCAAACACCGCAGCAACAAGCATCATCAGGACATCGAGATCACCGAGGCGGAACCTTCCAGGACGCCCGAGGATTCGGTGTG
CAGTGCCGAAGGACCCACGGATACCTATGCGGAAATCTCCGGCGTGTCCCACTCAATGCTCTCTCACTCGATGGTCTCCCACTCCGTGGTCTCGCATTCGAAGTAG
(SEQ ID NO: 182)

Start ATG: 1 (Reverse strand: CAT)

MVPSRGKWPCRMFYDYTDIAEDAARQFIEFLSGKFPNANTPIAIDEPTRAEVSRRANGIASYALNEDDNLLAFAIAAPSIHTVVVKFRDNVTIPPDQVHNKAYLGSYWRELG
AAWNSTDGTQEWGAPFRDCNLLTRRWLWPFRISFSEHRIKVVAAAFIAADEDVCNDGLEEVFGRRHGCDRNTTFCLLTENKPAATRDVYTCLCRESYYLPNSTLQGFRGDRV
ELSEGYDNYSCIPCPGGCTNCDSNGVCLTFQEEEVLNVDACLRLLVAIVLGACILCCIVLGVIVFRQRKCKAIASGMWTVLETILLGIVLLYASVAVHFFPASTERCLLEPW
LRELGFITCYGAIILKLYRHLVDFRTRKAHRWVLRDVDLLKYLGTMVFAVICYMAAFTASSLDLLESAQLESLREADTNTCHPLKWELVTQTSEMLILCFGLHLSIASRNAN
TQFRERQFLVTALTLEFLVSSSFYFLRFVYLPEMSPSAILLALFIRSQLTNSFALGLIFVPKLWYQHKQGTSHDAGQRLGGGYAGLCLGDPDIGELTISEMSPEDIRAELKR
LYTQLEIMKNKTLRQDNPHISKRRGGRKAGHRRFSLQKKGSKDKALSAKHRSNKHHQDIEITEAEPSRTPEDSVCSAEGPTDTYAEISGVSHSMLSHSMVSHSVVSHSK*
(SEQ ID NO: 183)

Name: Metabotropic Glutamate-like RECEPTOR
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384249
CTCGTGACCTTTGAACCCGACGCTGGCACTTTGAAACGGAAACGGAAACACACACTCACACACACGTAATTTGGTTAGTCTGATATCGATGCACTATATGTTGGCTAATTTA
GACGGACTTTGGTTTTGACCAATCGCCAGGGTCCTACATATATGGTTTTGACTATATATTCTCGCTATGTTTAGTTGATTATACGATTTAGGGGTTTACTTTAGGTATTTTG
AGCATTTTCAATTCATGCGATTGTCGTTGAAATAACAATTGGGTTTTAGCTTTGCTTTAGCTGCCGTGACTTTTGTTTGTATCTCTATAAATTTATCAATATTCTCCGATTC
TAATTGCGTTGGGGGTGTGGTTGTTTATTCGCAGGACTTCTCACGTAGTTGTCTATAATTTGCGGTCTCTTAATTGGCACAGTTTTAGTTTTAGCTTTAGCTCGTCGATTTT
GCGAATCCCTATGTACACGCATAGCTCAGTGAGTTGCAGTGATTGCCATTTTGAATAATTTATTTGTGGTTATTTCTTTTCGTGAGCAAAATGCTTTACGGCTTACAGATTT
TCGATTGGCGTACGTACATACATATTACATATTCATATACATTTATCGAAATGAGTTTACAACCTGTGCTTAATTTAGTCGTACAAAAAATACTCAGTATTTACAACAAGATA
TAAACGACATAAAATCATAAATAAATAAGTTATCCTATACAAAGTCCCGATCTGTGCTCCAAGGAACCCGGCAAACTCGAATTGAGATTAAATGTTTTAAAAACGAAAACAA
AAATGAAAAAATGTGAAGAAAATGAAATGGCACGCTCCACAGATTGAATAATGTCAACATTACGAGTAAATCAGCCGACGCATTTGCCGGCATATTAATTTTACAAATTAAT
GGAACACTTTAATTGCATTAAATTGATTCCACCCATCCGGCTGGCATCGATTGATTCGATTCGCATCCTTACACCAGTCGCACCATTTGCACCATTTGCACCATTTACACAT
GACTCGAGGCCGTCGAGGCGGCGGCATTGTTGTTCTTGTTCCATTTGTTGATGTACTCCATGGGTATTTCAATCTGGATCTCCTGCACATCGTCGCCCGCCAAGCAATCGAC
ATTGGATTTGCCCTTGCCCGCATCCTTGCCCGTATCCTTTGCATCCTTCGACTCCCTGCCATCCTTGGCGGCAGCTGCCAATTGTCCGCCCAGGCGGGATCCGCCCGCCGCC
AGCCGCTGCTGCTGCTGGGCGATGGGGGTCAGCAGCGCCTCATCCCTGTATTATAAGTTGTCCACCATCTGAGACTGGCGGCGATGCGAGGGCGTTGTCGAGTTAAGGTAGC
TCTCCGAGTAGACGCTCTTCGCCCGCTGGTCGTAGCGCAGGCTGGAGCGCCGGATGTCGATGTCCAGAGGCAGGTGGCGATCCTTCCACCAATTGCAGAACAGGCAGAGCAG
CGTGTTGCGGAACGCCTCCCGGAAGTCGCGGTTGAAGTACGCGTAGATCAGCGGGTTTAGCGTTGAGTTGAAGTACCCGATCCAGAAGAGGATTGAGACGACTATGTCCGGC
ACTTGGCACTCCTCGCAGGTCATGGAGAGTGTGTACCTAAAATTAAAACAATATAATTATTAGCTTACACTCATTATCTTTAACTTAATTTAAGTCAATTTTAAATTAGTTG
CACAGGTAATGCAATTTTATGTACTTTTTTTTTCCCGGCATTGCATAATATCTATTTTTTAATAATACAAAAGCAAATATAAAAGATCAAATGAATATTTTTGTTTCATAAA
AGTTAAACCTCATAAACGATCTATATTTGAATACTTTGATCACCATTTGAAATCTTTCTGCATCACTTTTTGTCTGTGTAGCACATAAAGATGAGATATATATACCCATAAA
TTCCAGATGCATGAGTGGGCCAGGGCTTTGCATAAATGAGATTATACTCTGGGTGGTGAAAAACAACAACCGCCGAAACAATGGCGCACGGGCCGTATGAGTAATATTA
ATTATGCATGCTTGACATTAGATAAATTATGCCTGGAATCCCATGTCAGCGGGATTGGCCTCCAACTCTCAACCAAACGTTCAACCTTCGACGGCTGGTTGGTACTCACCAC
AGGAAGAAAGGCAGCCAGCAGAGGATGAAGGTGCCCATGATGATGCCCAGCGTGCGTGCGGCCTTGTGCTCCCGCTTCATTTTGATTAAGTGCTTGTCCTTAGTGGGGGTGT
GCTCCTGCTCGACCTCGTGCAGCGTCAATGTTTTCGACGACCCGGATCCGCTCAGCGCCTCGCCTGCCGTAGGAGATAAGATGGAGATTTGATAAAAGTCCAGGCTGGAGAT
GGGCCACTGAACTCAATGCACTCGCACTGCATCACTGTCCAGGATGCTTGTAGAACAAACTAGTGATCTTTCACGCTCACTACCTGTTTGCTATTCAGAAAGTAATAAGACT
TTCTGTCCTGGTAAATTTATATTTTGAAGTATTATTTCCACTTAAAACCATTATATATATATTCAAACAAACATTTCTTTACAGCCAAACTAGGTAAGCCTTGTTGTTTACA
ACCCAACTTTGACTCCAAACTCACCTGATGGCTGCATGGATGGTCGGTGCATCAGCATGGCATTGCCGTGCCGCATCATCAGCTGCTTCTCCTGCCGATTGGCTTCCCGGAA
GATGGCCAGGTAGGTGAATATCATAATGGTGCAGGGTATCCAGAACGATATGGAGCTCGAGATGACGGCGTAGTACTTGTTCACCACGAACGAGCACTGCGTCGGATTCTGG
ATGACGAACTGCTGGTGCTGCGGCGTGGTGTACCAGCCGATGAAGATGGGCAAGAAGGAGAGCAGTGCCGGCGATATCCATGTGTTTAGCAGCATAATGCCGACCACGCGTT
TCGTCATGCTAATCGGATACTTGAGGGGCTTAACAATAGCATAGTATCTGTGAAGTGGGCAAAACGCCAGGTTATTGGGCCCTCATATAGCTATAGTTCTGGCTGTGGCATG
CAACTTACCTATCCACAGATATGCAGCATAAATGCAAAATACTCGCTGTTGAGAGTAGACATCGAGGCTGTTCCACAAATCGCACAGGAAGGGGCTGAAGTTCCACCTGAA
AGGGTCCAATGAATGGAGGCATTAGTATCGCGATGCTTACCCAATAACGAGGCAATAAACTTTGGCCTTTACTTTCTAACGTCATTGTATCTGATTATCGAGTCCATGAATT
TGCAGTATATATTTTTGTTGACTCATTTCTGGTTTTTTGCCTCACCTAAAATGCCACGCGAAATACTTTTATGAAGTCATCGCTTTAAAACTGATTTCCTGGTTTTTAGCTG
TGCAAAAAATTTAGTTTTTCACGATTAATTCCCTTTGTGACAACACAGTTATTTACGTACTCTGATTAAAATCATTTGATTGTGGGTAAAATTGCGCTAGTTCTAATATTTA
CTACCATTTAAAATACAGATATATAGGCCATAAAAACTCACTAGATATGCGCAATAAGGTTGCTCTAAATTTAAACGGCATTTTACTGCCAATTCCTGAATGTAGGTCTATA
AAAAACTCAAACTGACTAGACACACTCTCTTTTTCGCCTATTTTCGTTATCAGGCCGCCTGATTTATGTGTTTATCGATTTCATCTTAGTTTCGGTTGCCAGGTGACATTGC
CCCTGAACTGCCATTCTGGAGGGCCAAATTAAAAGCTTAGATGCACTGTGAGACCAGAAAATAGGCCATAAGGCCGAGGTGAAAAGTGGCAAGGAAAGCGGAGGCGACTCT
GTTCGCTGCATGAAAAACTCATTAAAACTTGCGTCGTCGCTGTTCGCCTGTCGCTTCTGCCGTCGCCCTTGGTTGCCACTCCGCATTTGGCCATTTCGGTGAAGGCGGATA
CCCGGCCTTTTTGGGGCCAGGAGCCAGTGGAATTTATTATCGGGCGCCCCGTTTGGCATCAGTTTCAAAAACTTTTCGACTATCCAACGAGGCTGTTTTTGCTCGCTGGTTT
TTGGTGCTTTACACTACATTGCCATTTGGGGCGGGCGCGCCATCGGTTTTGACACTTAAGCCATGAAATGGCAGTTCCTCGTGCTGCAATTATGCAACTGGCAACTGTGAAG
GTAAAATGTCTAATGAATGCCAAGTGGCGGCGAGTTGGCTGTTGCCAGAACCAAACCGATGCCAATTTATATGCAACTCGCGCTTGGCCAAAGTTTTTC
(SEQ ID NO: 184)

Exon: 3243..3209

Exon: 3131..3033
Exon: 2959..2601
Exon: 2303..2126
Exon: 1604..1298
Exon: 1246..1001
Start ATG: 3243 (Reverse strand: CAT)

Transcript No. : CT21650
ATGGACTCGATAATCAGATACAATGACGTTAGAAAGTGGAACTTCAGCCCCTTCCTGTGCGATTTGTGGAACAGCCTCGATGTCTACTTCTCAACAGCGAGTATTTTGCATT
TATGCTGCATATCTGTGGATAGATACTATGCTATTGTTAAGCCCCTCAAGTATCCGATTAGCATGACGAAACGCGTGGTCGGCATTATGCTGCTAAACACATGGATATCGCC
GGCACTGCTCTCCTTCTTGCCCATCTTCATCGGCTGGTACACCACGCCGCAGCACCAGCAGTTCGTCATCCAGAATCCGACGCAGTGCTCGTTCGTGGTGAACAAGTACTAC
GCCGTCATCTCGAGCTCCATATCGTTCTGGATACCCTGCACCATTATGATATTCACCTACCTGGCCATCTTCCGGGAAGCCAATCGGCAGGAGAAGCAGCTGATGATGCGGC
ACGGCAATGCCATGCTGATGCACCGACCATCCATGCAGCCATCAGGCGAGGCGCTGAGCGGATCCGGGTCGTCGAAAACATTGACGCTGCACGAGGTCGAGCAGGAGCACAC
CCCCACTAAGGACAAGCACTTAATCAAAATGAAGCGGGAGCACAAGGCCGCACGCACGCTGGGCATCATCATGGGCACCTTCATCCTCTGCTGGCTGCCTTTCTTCCTGTGG
TACACACTCTCCATGACCTGCGAGGAGTGCCAAGTGCCGGACATAGTCGTCTCAATCCTCTTCTGGATCGGGTACTTCAACTCAACGCTAAACCCGCTGATCTACGCGTACT
TCAACCGCGACTTCCGGGAGGCGTTCCGCAACACGCTGCTCTGCCTGTTCTGCAATTGGTGGAAGGATCGCCACCTGCCTCTGGACATCGACATCCGGCGCTCCAGCCTGCG
CTACGACCAGCGGGCGAAGAGCGTCTACTCGGAGAGCTACCTTAACTCGACAACGCCCTCGCATCGCCGCCAGTCTCAGATGCAGCAGCAGCGGCTGGCGGCGGGCGGATCC
CGCCTGGGCGGACAATTGGCAGCTGCCGCCAAGGATGGCAGGGAGTCGAAGGATGCAAAGGATACGGGCAAGGATGCGGGCAAGGGCAAATCCAATGTCGATTGCTTGGCGG
GCGACGATGTGCAGGAGATCCAGATTGAAATACCCATGGAGTACATCAACAAATGGAACAAGAACAACAATGCCGCCGCCTCGACGGCCTCGAGTCATGTGTAA
(SEQ ID NO: 185)

Start ATG: 1 (Reverse strand: CAT)

MDSIIRYNDVRKWNFSPFLCDLWNSLDVYFSTASILHLCCISVDRYYAIVKPLKYPISMTKRVVGIMLLNTWISPALLSFLPIFIGWYTTPQHQQFVIQNPTQCSFVVNKYY
AVISSSISFWIPCTIMIFTYLAIFREANRQEKQLMMRHGNAMLMHRPSMQPSGEALSGSGSSKTLTLHEVEQEHTPTKDKHLIKMKREHKAARTLGIIMGTFILCWLPFFLW
YTLSMTCEECQVPDIVVSILFWIGYFNSTLNPLIYAYFNRDFREAFRNTLLCLFCNWWKDRHLPLDIDIRRSSLRYDQRAKSVYSESYLNSTTPSHRRQSQMQQQRLAAGGS
RLGGQLAAAAKDGRESKDAKDTGKDAGKGKSNVDCLAGDDVQEIQIEIPMEYINKWNKNNNAAASTASSHV*
(SEQ ID NO: 186)

Name: Beta 2 Adrenergic Receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384249
GTGAATCGTTTTTGCAGCGAATGGCGAAATTTCCTACAGCTTGCCAATTGTTTTGGTTGTTCGCAGTTCATTCTTTTTACTCATTGGTAAGATTGTCTTTCTGCATGAAAAT
TCATAACGTCAGCTACACAATTCTCTTCCATTCAATATTTTGTCCATACATAATTTCGATTGCGTAATTTGTAATTTTCAATAACTCTGTTATGTAGATTTGGTTTGCTAGT
TTTCTAGGTTACATATGTATGATACTTAGAGCTAGAGGTAACATTATGGGATTGCACTTTTTCCATGATTTAGGCTAGTTCTATCTGTAACAACATTAAAAAGTGTGTTTAA
TATGAAATAAAAGCAGGGAGTTCATTCCGCTCAGGATTTGCAATCCTAATATACGATGATGACACATGTGAATAATCCAATTTTGGATCGTATTAATACTCCAAAAATCGAT
GGAAAAAGTGTTTCGCCTTCGACAAGCCGAAGTGGTGTACATAATTCTTAGATGATTCATTTGACAACATTTCACTTGAAGTTGGAACTTTTCTACGCAAAGAGTCGCTCGA
TTTCTAATTGGCTTTCGCTCTGCCATATATGTATAGTTATTACGTATATCTTCTGTTTAGTTGGACAAACAACACTTATATATGGATAGTATAGTTCCGGGGCAGCATTGTC
CCTGCAGCCGCGGCTTGGCGGGCTTAGAACACAACTTTAATTTAAATTCCCTATCGTTGCTTAAACGAAATTGTACACGTATACATAAACACACATCGAGTATTGCTTAGCT
TCAAGGATTATGCTAGCTATATCCTGGCACAAATGCCACTGGTATGGCCGTTGCTCTTGTTACTTCTCTTCACTTCAGCTGGTGTTTTCCCTCCTTGGGCACATATACATGT
GCGACGATTCACTCAAGTATTAAATATCACTTATAGGGGCATTGTTCCGCATCCACGAGGGTGGACGAAGCAGGAATGTGGCCGGGATGCAGACGCTATCCGGTCTAGACGT
AGTAGGCGTTGTACGGATTTCGTTTCTCCAGACAGGGCAGCACACACTCCAGCGTATTGCGGAATGCCTCGCGAAAATCGCGATTAAAGTAGGCGTATATAAGCGGATTTAG
CGTGGAGTTGAAGTAACCGATCCAGAATAACACCACCACGAGCACATCGGGACATGGGCAGGCCGGACCGCAGAGCGATGTGATGACATACCTGGAAGGAGGAGTGAGAATC
ACATGAATATGAGTCACCAGAAAAAGAAGAATAAGAAGGCCAACAATGGATTACGGATGGCGGTACCACATCCGTACGGCCTATCTACCGTCACTCACCACAGAAAGAAGGG
CAGCCAGCAGAGCAGAAAGACGCCCATGATGATGCCCAGGGTGCGGGCGGCCTTGTGTTCGGCCTTCCATCCTTTGGCCGGTCGCACGGAACCGCCTGTGGGGCAGATAGAG
AGTTTAGAAGGGATATGAAAGCTGGCAATTGGCTGATTAAGTTCAATTCGAATCGATTCGTTTCGATCCGCCAGTCTGTTTTTTTTCTTTTGTTGGGTCATAGCAATCGACA
GTTTTCTGGTTCATTTCTGTCTGTCTGCGAAAGCACTCGAAAAATGGGTTTCTTTTTAAGTAGTCTATAATATGCTTAATGTGGTATTTGTTGTCTACCTGTTGGATGTGTT
TGTTTTATGGTTTATTGCTTCTAGGATGAAATTGTTTATACAGCTAATACTGGAAAGTATAGGTGTTTGCTTAAGCTTTAATCGGCGATACCCTTTTATGTTAACTATGTAT
GTAAATTAATTGAACTTCAATCAATTAAAATACAGAAAATTTCAAACAAACAAATTTTAAAGAGGTAACTAAACTAATCTTAAAAATTATTAAATATTTTCTCTGTGTGTAA
ATTGCCAAGAAATGCGGTTTATGTTCGTTTTATTGTGAGAGAAAGATTATGCCGGCATGTTGCGTATACGCCATGTGTGTGGTATCCTTGTTTAACTTTGACATTAGTAATA
CTTTTGCTGTTATATGTTGTTGGCTATCGATGGCATTTCAAACCCCACACCGCCCCGCTCCCTGCGTCCACCTTGCAAGTGACTGTCATTAAAGTGCTGGAAAGGCCAACAT
GTGTGCAGTATCTTGGTGCTCGGGCGTGTGGGTGTGCGGGAAAAGCAATAAAAACGAGGCAAACAATTGAAAATATTTACACAGGACACTTGTTATCACAAAGACATACAAT
CCCCCGCATCCATCCAGTCCAGTGCGCCCGCCCCAAACATAAACATAAACATCACTGGCGGCCAGGAGGAAAAACAAATCAGACAGACAACAGTGAGCAACAGGAAGCAACG
GACCATCGCCCTTGTTCACCCATTGCCCATCGCCCATTGCCCATGTTCCCCCATCCAGTTTGCCACCTGGATCCCTTGATGGACGCCACTCGGTGATGGCCACAGCCAGGTG
AAATCATAATGCAAAGGCTGTGTGGAGGGAGGCGTAATAAAGCAAAGCAAACGCACGTGCAAACGGAAGCAATTATGTACAGCATGAGAGCAGCGAAAGGTGAAGGATTCGG
ATACGGATTCGGAGGCAGATTAGGTGGAGGTGGAGACAGCATGAGGTTTCGATGGTGTAACACTCCGGATTCAGTTGCCTAAGCTGGTTAGCTTGGCATCCGGATAAGGAGG
TGCCAAAAAACTGGCACTTCTGTGGGTATTTGCCACTAAAGCTGGGGGCTACAAAAATTTGAGGTAGTTCTTTTCGGTTTCATTTCACGGTTGATTTAATGATACTTAGGTT
AACCACCGTTGAAATGATTGGGATTGCAGAAATTCAGGGATATCTATTCTAGTATGATTTTTGGTAAGAACGTTAACTAGTTTCAGGTGTATTCAAAAGCTTTCAATTTTAA
ACATAATAGGTTTAAACATATACGGAACAGTAATATTGATTGTTTTTTGACTGACAGGATATTTCATATGAAATAAATAGATTTTTGATATGTTTGGTCTAATGCAGTGGCA
TAAAATACGAAGTGAATGGGACTTGATTTGATTAAGATTTTCTGATGATTATATATGTATTGATATAGAGAAACTATAACTATTGGTGTTGTGAAATTAGTGAATTCTGAGT
GTGAAATGTATTATAATATAATGCACAAGTTCTGGACTGTGGGAAATCGAATGATTCAGATGGATAATGGGTAGCAATGAGATAGGTGCTGAAGCTGAGACAAGCAGAAAGA
CAGAGATAGAGAGAGAAGCGGGAGAGAAAGAGCGAAACGTAGAGACTGGGGGAGGAGAGGAGGGGAGATACGACAAGACTCACCGGTTGCACCACAGTGACTGCATCCGGTT
TTGAAGCGGCGCTTCCGGTTCTTGCTCTGCTGCCAACATGCTGTGACCGCCGCCGTTCGCTTGTGCCAGAATCGAGTGGAACTGGACGAGGAGGCGGGCCTCTGGCCAAGAT
CTGGAGTTTGCTCGTGAATACTACGTTTACTACTAGTGGTTACTGTGGTGTCGATTGTGCTGTCGGCGTTGGCGGTTGCGGTTGCGTTTGCGGTGGTGCCGGTGGTGCTGCT
GATGTCGGCGCCGCAGAGCGGACACTTGGGCAGCGGACTGGGCACCGCGGCGAAGGTGATGCCGCGGGGCCGCTGCTCAATGATCTCCGCCTCCTCCTCATCCACAAAGATG
CAGATGCCGCCGCGTCGGATCAGCGAGGATTTATTGGAAAATTCGTGCTCTGCTTGCAGTTGATTGTTGCATTTTCGTACGATTTCGATTGTCAATTTGGCCATGGCATTTT
TGGGGTTTTATTGGTTGGATTTGGTTTTGGAGTAGAAGTAGTTTGTTTGTTTGGTTGGACATAATTTGGTGATTTCGATTTACATATATACACATACAACACACAACGAAAT
GGCGTGGAAGAGAGATAGATATAAAGAGATAGAGACAAGTAGTAAACAATGTACACAGGGTATTTCACAGGATACGAAAGGACTTCAACTTCTGAGATCACATGCAGCCATG
CACCGGCAAACGCAAAGGATCGGGATCGAAAAGGGAATGGATTCAGATTCATTATCGTTGGCTCCCTGTCTCTGGAAGTAGGTCGTCCCTTTCTGTCTTCGTTTGCAGGTGC
AACTAACTAAATGCTATACATTATATTATATATATATAGTAAGCTAGTTAACTAAAGTTTCATTCGAAGCATTTGAGTATTACTTTTCCAATGAGGGTTTATAGGTTCTATA
CATTTTAAAACACTGTATGATTTTTCAGAAATTCGCACCGCTTTTGAAAGGAACTTAAATACACAATTCACAGTTTTCTTCAAGGTAAGAACTATATTTAAATATGGTAAAA
GTTCGTATCTTTATTTTAAGATCAAAATTCTAGTAAGACCAGATTCCTTATGTTAAACCCGAATTCAGGCCAATTACTCCTCAAAAAAACAAGGACAAAAGAATGTGCGAAA
AAGTTAAGTCAAACATTGAAACACTTGAACTTTTCGTGGAAAACTGACAAAAACAGAACGAATTTCTCTGGGGCAGATATAAATTGTCTTAAAAGAGCTACTTAAATCAAGC
ACGAGCTATACAGTCTACACAATGACCAAATGGATGACTACTCACTTGCGGAACTGGTGGAAACGCGGCCCATCGGGATGTAGCTGCGACTGTCGCTGTTCTCCTGCGAGAA

GTTGATGACATTCACGTCGTAGGTCACCACGGACGCCTGTCTCTTCCGCACAGTTTGGATGCCATCGTAGCCTGCCAGGATGCTGCAGCGGCCTTGGCGGCCACCCACTCCT
CCAACGCCACCAGCCCCTCCATTGCCGGAAACGCTCACGTCCGCATCGCCGGCGCTATCTTTGCGCTTGAACAGTGGCGACGTGACATCCCGGTGCTCATCGCCGCATCCCT
CAACCGGCGACAGCGTGGAATCGAGTATGATCTCCGGCTTGGGCGATATGTTGGACTGTGAGGCCATTTGGGGATTTGGCTGCGGCACTGGAACGGGCATAGGCAGGCTGGT
CACCACATCCAAGGACACAGTGGATGTGTCACTGGGCGTATAACCAGGTGCGGCGGCATGCGGCGGCGGCGATAGGAAGTACACATTGATGGACGCCGACTGCGCGTTGGAA
GAGGCGGAAGAGTCCATGGCACCTCCGTTCGGCACGGAAATGGCCGGCGGAGCCGGTGTGTTCGGTGGGGACATCAGATAGTTGGCGCACACCTGGTCACTGAGCGTGCTCA
GGCGCTGTGGAGCACCAGCCGGTCGGCCCAGCGGAAAGCTCAGCGAGCGAAAGATCTCGGCCAGTTCCGAGCTGCTGGCAATGAGCAGGCCACTCGGTTCTGGCAATTCTGG
CGGCGATAACGATGGTGGCGGCGTGGCCGGCTGCTGGGGCGACTCCCCAAAGCAATCGGCCAGCTGCGCCAACTCGTCGTCGCTCCGAATCCTCTCGGGCGGTGTTTTCAGC
TCACTGAAGTCAATCAGCTGTACCCCGTCGCTGAACTGCTGGATCACACTGGCTCCACTCTGCTCGCTGAGCACATTGATGTCCAGCAGGATGTCCGGCTTGTTGGGCGTGG
AGTCCGTATCGCTGTGCTGCCGCTGGCGTCTCTCAAGTTTGCGTTCCTTGCCGCCACCGCATCCGTTGTAGCTGTGCGACAATTTCCGGTTCCGGGAGCGCCGGAAGAAGCT
GGATCCCTCGGAGTTGGCCTGATGTCGCTTCTTTTCCAGCGTAAAGCACTCGGTGCTCGAGCGTTTCGTCTTCGACAGCAGGGACAGAATGTTGAAGCCCTTCTCCTTCCCG
GGCACTGGCTGCTCCGGCAACGCCCGCGATTCACTAATGGTGCGCAGGAACTCCGAATCGCTGGAGGCGAATGCCTTACGCCCACAGGCCGCTGCGCCCGCCTTCAGCCTGA
TGAAGTCCGTTTCACTGAGGGCAGCGTACTTTCCTCCGCCTGGGCCACTGATAATGGTGCCGCCCGTCTGGCGCTTAAAGTACAGATAGTTGTCCTCTATCAGGCGCTTCAG
CTCCTTGTTGTCCTCGGCCAGCGGCAGGTGGGTGTCGTACAGGTCGTTGCCCTCGCTCTCGTCGCTGTTGAGACCACCAAAGACCCGGTACGCAGTGGACTCCGGTTCCGGT
TCGCTGTTCCCACTGCCGCTGGTCAGGCTGAGCTGCTGTTGCGGGCCATTGGGATGAGAATGCTCCTTATCGAGCTCCTTCGAAGGGATCATCTTGCCCTCCGTCCATATCT
GCGGATTGAACACTGGCGCCGGTGGCTGCAACTGATTGTGGGAGGGCTGCTGTTGCTGCAGTGCCATCATCGAGGGAGCACTGTCCGTGTGGGTGACCTTCTCGTACCGCTC
CTGCTCCAAGTCCCTCAGATCGATGCTGCTGCGGCTGAGGCGGCGTGGCGGCGGCGATGGCACTCTCAGTTCATCGCACTCGTCCCTATCGTCGTCCTCGTCGGTGGCCGGC
TGCAGCATGTCCTGTGGGCAAAGGACAATCACTCAAAGGGAAGTCATGGTGGCGCATCAATTACGCACAGCCACAAGGGAAAAGGAAAAACTCTCAAGCGGAGTGCTTCCCT
CTAGCACCTAATCTAATCGCTCCACATCCGTTCATTGAGGATGGGCAAAAAGCTTTAGAGCAGGATGTCAATGTTGCCCGTTTGTATACATGAACTGCATTGCTTTAAAGTC
TTTGTATATTTATTTAGGATACTATCTGGATGTATCAGCAACTTTACCAAATTTTACTCAATAAAAAATAAATAATGTTTTATTCAACTACTAAGAGATCTTCAGAATGGCA
AGCAAATGCGAGTACTAAAGCATTTATCGGTTTAATAAAAATAAAACATTTTTAAGTATCTCTAGGTATTTTAGGCATCTTTGGATTAATGTGTATTTTTTCATTTTGAACA
TCTTTGAAATGTTTATTGTTTATTGAATGATTTATTGAATTTATATCAATAAAATGCAGCAAAATACTCTTTTGGCTTTCAATAAATGACATCTGTATCTTTTCGGTTCAGT
AACATTTAGTAACCTTCCAAACTCACCTCCAAGTTACTAAGCGCACTGTGCGTCTCCTCCCGGGCGGATGTGGCATTCAGATCACAGTCACTGGGATGCAGATAGCTCAGGC
TGGTGGGCTGCTGGGTGTGGCCCATGTGGACGCTGTTAAGGAGGATGTTGGAGCTGGTGCGGCTGAGGGCCTTGCGTTGACGAATCGCCTCCCTGGAAAACCACAGAGCACA
GATCAGCAATTAAGCAAACTCATCCACTACGACCAGCGACCATCGACCATCGACCATCCGGGTAAGAACTCACTTAAAGATGCGCCAGTACATCACCAGCATCACAATGCCG
GGTATCCAGAAGCTCACCGAACTGGAGATGAGGGCGTAGGCCTTGTTGACCACAAACGAGCACTGGTCCGGATGCAGGGAAATCTCCCGCAGGTGCTCCTCCGTCGTGTACC
AGCCCAGAAAGATGGGCGTGAAGGATATGAGGGCGGGCAGGATCCACACATTGGCGAGCATGAAGCAGACCGTTTTGTGTGTCATATTCAATGGATACTCCAGTGGACGCAC
AATGGCGTAGTATCTGTTAAAAAGAAGAGGGGGTCATCTTTATTAATGGAAATGAATTGCGGACAGGTCCTCCGCACAAAGGTGCCAGGTGGCTTTCAGCCATTTCCTCCTT
CTCCTCCTCCTCCTTTTGGCTTACTAATTGATTATCGCTCTTTGGTTGGCTGAATAATTGTGTTTCGGTACTCAGGTTGCCATGCACAGAAATCGATTTATTAACCTTTTAA
AGATCATGTTGCATTTAAATGAAGGATAAATGCTCTTAACCGAATTCATTATCGGAATCCGTGAAAATAAGGTGGTGTATTCTTTTTAATTATCAATATATTAATAAATGCA
ATTGAAATGAAATAATTATTAATCCACGTATATTGTACCTCATTATTTTATCGAAACTTAATGCTCCTTTTAAAGAGTAATGATTATGTTTGGAACATTACAGATTTATTAG
TTTAGCCAAATTTCCGTTAGTGGCCATCATTAATTCTCTAATTTTGATAGGTTCTGGTTGAATTTCCCTTTGTTTCTTGGTATGTTTTGTGGCCCGTCGAAAGGTGAGTGAA
ACACAAATTAATTAAAACAGACAAAAGGCCCATGACTCCAGTGTCCATCGACCATCCCTCCATTGTTGCTGGTCAGCCCGAAAATCGCAGCTGTTGGTGCCTTCCAAGTGGG
GAATGGCGCCAATTAACGGCTTGACAGGATGGAGAAGGTTGTGGGAGTTGGGTGGAATATGGGTGGAATATGGGCAACGCCCACGTTTCGGCCAGCTGCAAATCATTTGGGC
CGCATTCATCATACGGGCGGGCCTAATGTTTGCTTTTGATGTGCGCGCCTTTCATAGTAATTAATTGTTTCTTCTTGGCAATTCGCCGGGCTTTGGTATTGAAAGTGATTTG
GTTGCCAGGCATCTGGGCAAGATTAATGGTATTTGGCCAAACTAAACGGCTTTCTTCAGTACCATAGCCCAGCACAATTGACTTTAAACCGCTTAATATATAGATAGAGAGC
GACATAAATGGGCACTCACCTGTCCACTGATATGCAGCACAGGTGCAATATGCTGGCCGTGGAAAAGTAAACATCCAGGCTGTTGTACACGTTGCACATGAACGGTCCGAAC
ATCCACTTTCCGCCGGACAGTTCCACGGAGGCGTTGAATGTCATCGCACAGAGGGCCACCAGCATGTCGGCCATCGCCAGCGACACAACAAAGTAATTGGTTATCACCCTGC
GAACGAAACAAAAGCCCAGGAATTTGTGAATATGGCGTGGTTCTGGGCGGGATTTGACCAGCTTCTCCCCACATCTAACCACTTGTCACAGCGCAAAACCACGCCAAACGTA
ATGGCATATGGCATGCATCATATGTGCACATGCCGCTGTACTAAGCGCTTTGGTTGGCTCCTTCCGTTGGCCAAGCAGTGCACCACTGCTTTCCATTTCCTGTTCCGGTGTT
CAGGAGTTTGGAGGCAATAGGTCACCGGTATGCAGCAATTTCCTTGGCTGTCATAAGTCACTCGACTCACCAGGCAGCTGCCACCTTCAGCTGCAAGCTACTGGCTTGTGAA
GTTATCATCTTGTCTTCGGGGCGTATGTGGTTTAAGGAGCTGTTTATGCGAAGATGAAAGCACTTTTGTGCTTCCAAGTCATTTATTACATATTGCGACATCTATTACGTTG
GAAAGTGTAGTTTTACAGGGGCTGACAACATTGTCTAAAACATGTTATTAACGTATATTGTAATATTGTAGTAATGAAGCTCAATTAAATCAAGATTATTAATCTGTGCATT
TAGAAGGGCTGTTAATTTCGAAAGATCCTTTTATGTCAAGCCGTTTTTTCTTTTATGCACTGCTTTATTTTGTAATATCTCAACGCAATTCTATATGGCGAACGTTCTATCG
TAAGGTAGTAAATGATAATCTGAGCACTTTAAGCTCCTTTAGCGGTATAATAAAGACTTTAATCAACATCTACCCTTGCTGCAGAAGGCATAAACGTTGCGCAACATTAAAC
GGATGTCGTGGGTGTACGTGTGGGGGCGTGGCCGGGGGCGCTGGCGGTGGCAGAGCGCTCAATGGTATGCTACACTTATTTGCACTGCACTCTGGGGCGTCCTGTCATCGCG
ACGCTGACATGGCAGCGATGGGCTTGGGATCGGCTTGTGTGCCGCATTTAACTTCGTTTTGTTCTTTATTATTCCACTTGTCGCTCATTTAAATTACATTTATCGCGATGGC
GAAAAGGATTCGCCCTGCCGCGTCACTGTTTGTTTGCCATTTCCATTGTCCTGCGACCCTTGCGGCAATGACCCTCAATTAAACTCAGTTACAGGTGGTGAAAACAAAAGTA
ATTGGATATAGACTTACCGCAGCTTTCGATTGCGCTGCACTGAAATGATGACCAATGCATTGCCGAGGACAGCGGCCAGGATAATTGACGAGAAGATGAATCCTTTGAGCAG
CAGCAGCGATAGATCCAGCCAGCTCTCCTCCAGCTCCGCCTGGTTATCCACCGCAATGTCACTGCCATTGGCCGTTATATTTGGGCCGACTGTCGAGAGGAGCTCAAACGTG
GAATCCGACGACAAGTAGGATGAAAAGGGCGAGGAGTAGGAGGCATCTGACGATGTGGCTGTTAATGATAGCGAAGTCAACGAGGCATCCACAAATTGCGATGTTATAGGCA
AACTGCTCGTCGGCGTCGTCGTAGTTTTGGCTGCTGCTGTCGTGGAAATGTGGCCTGTTAATGTTGCAGGTTGCAGGTGGCTGGCGTTGGTGGCTGATGTTGCTGCTGCTTG
TGATGTTGCTGCTCCTGCTGCGGCTGTTATGGGTAATGTCATTGTCGTGGCCAGAAGGTCAGCCACGTTAACCCCTGACATCTCGTCGATTTCGCAGCCCAATTGAGTCGCA
GCTCATTCAAGTTGCTCCGCCAATGCTTTTTCGCTGTACTTGTCGTTTATCTTTTTCCCGGCTGCTGTGCGTAATTTAAAACTCGAGTAATTGCAATCTGCAATGCAAACAA
GGCTGCCGAAGTTTGCCGTCATCCTGGTTGCCAATGTTGCTCCTCTGCGATTCCAGTATTTGCTCACAGCTCCGACTGAAGGTCTCTGCTCCATTTCCCAATCCTCCATAGG
ACGCTTTTCCCGTTTGAGTTGCTGGAAATTGGCAAAAAGTCGAGACTCGCGCTTCGATTTGGCACTCAAGCGCGCAACAGTCGGCGACACGGGAACTGCAAATGAATAAATA
AATAATTTATGAAATAATCGGCGAGTGTGAGGCTGTGAAACTGCATGGCAAAATTGGTAAAATGGGAATTCGAACGGTTGCGCCGCCCGGGGTGTTTTTATGTACATACGCA
ATATTCAACAGGGGGCATAAACCCTATTGCCTATTTATGTGCCGCAGTAACACATAAATGGGATTCTTCTGTTTCTCCAGTGCCCCGTCAAATGTCCGTTCGTTAAAAGTGC
GCCACTTCAACTTTGCTACACTAGATTAAAATAACAGAAATATATTGCGTATACGCAGTGTGTGCACTCGGCATGCATCTGCATTTCTATATCCCCAGCGAGACAGCGGTGC
TGTCAAGTGCTGACACTTGTGCAGAGTGGAATGTGGAATATGGACAAGTTGTTGCTTTAAAAGTGAGATTAAAGCAAGATTATCAGTGGGAAAGTTGTTTCAGTAGTGCCTC
ACGCCTGGGCGCAGTGAAAACTGCAAATCCTTTTGTCCCGGTACTTGTAAGTAAAACAAAAGACTCTTCATAATCTGTGGTAATCTCGAATTTTTTCTTTTTTATTATTTT
AAACTTTTTGCGACTCAATTTGGCCGTAAAAATGCTAACACATCATCAGGGAAAAGCTGAAAGCGTGCTTACACTTTACTTTCTCCCCCGCAACATTGAATTGAAGCTGAAT
CCGTCCTTGGCACTCAGACAAATCGCTGGCTACACGTTTTCATTGCACTTATTCTGTGCACTCTAGATGCCCTGCCCTCATTCCAATCTTCAGAGTTTTGAATCTAATATTT
TTAAGGATTTATGTACTTATACAGATAACCTTGCCGAATTGATTGCATTGCGTTGTCCTTGAGCTGTCGATTCGACTTGTTACTCAGACCATCCGGTCACATCCTCCTGCT
AGATTGGCAACTTAATGTTTTAATGCTCGGACTTAAACCACATCAGTGCGCTCTATCATCCTTGGCTTATACATTTCACAAGCCATT
(SEQ ID NO: 187)

Exon: 10847..10568
Exon: 10497..9986
Exon: 8844..8644
Exon: 7629..7354
Exon: 7260..7083
Exon: 3656..3332
Exon: 1439..1331
Exon: 1211..1001
Start ATG: 10497 (Reverse strand: CAT)

Transcript No. : CT21843
TTCCCGTGTCGCCGACTGTTGCGCGCTTGAGTGCCAAATCGAAGCGCGAGTCTCGACTTTTTGCCAATTTCCAGCAACTCAAACGGAAAAGCGTCCTATGGAGGATTGGGA
AATGGAGCAGAGACCTTCAGTCGGAGCTGTGAGCAAATACTGGAATCGCAGAGGAGCAACATTGGCAACCAGGATGACGGCAAACTTCGGCAGCCTTGTTTGCATTGCAGAT
TGCAATTACTCGAGTTTTAAATTACGCACAGCAGCCGGGAAAAAGATAAACGACAAATGTCAGGGGTTAACGTGGCTGACCTTCTGGCCACGACAATGACATTACCCATAAC
AGCCGCAGCAGGAGCAGCAACATCACAAGCAGCAGCAACATCAGCCACCAACGCCAGCCACCTGCAACCTGCAACATTAACAGGCCACATTTCCACGACAGCAGCAGCCAAA
ACTACGACGACGCCGACGAGCAGTTTGCCTATAACATCGCAATTTGTGGATGCCTCGTTGACTTCGCTATCATTAACAGCCACATCGTCAGATGCCTCCTACTCCTCGCCCT
TTTCATCCTACTTGTCGTCGGATTCCACGTTTGAGCTCCTCTCGACAGTCGGCCCAAATATAACGGCCAATGGCAGTGACATTGCGGTGGATAACCAGGCGGAGCTGGAGGA
GAGCTGGCTGGATCTATCGCTGCTGCTGCTCAAAGGATTCATCTTCTCGTCAATTATCCTGGCCGCTGTCCTCGGCAATGCATTGGTCATCATTTCAGTGCAGCGCAATCGA
AAGCTGCGGGTGATAACCAATTACTTTGTTGTGTCGCTGGCGATGGCCGACATGCTGGTGGCCCTCTGTGCGATGACATTCAACGCCTCCGTGGAACTGTCCGGCGGAAAGT
GGATGTTCGGACCGTTCATGTGCAACGTGTACAACAGCCTGGATGTTTACTTTTCCACGGCCAGCATATTGCACCTGTGCTGCATATCAGTGGACAGATACTACGCCATTGT
GCGTCCACTGGAGTATCCATTGAATATGACACACAAAACGGTCTGCTTCATGCTCGCCAATGTGTGGATCCTGCCCGCCCTCATATCCTTCACGCCCATCTTTCTGGGCTGG
TACACGACGGAGGAGCACCTGCGGGAGATTTCCCTGCATCCGGACCAGTGCTCGTTTGTGGTCAACAAGGCCTACGCCCTCATCTCCAGTTCGGTGAGCTTCTGGATACCCG
GCATTGTGATGCTGGTGATGTACTGGCGCATCTTTAAGGAGGCGATTCGTCAACGCAAGGCCCTCAGCCGCACCAGCTCCAACATCCTCCTTAACAGCGTCCACATGGGCCA
CACCCAGCAGCCCACCAGCCTGAGCTATCTGCATCCCAGTGACTGTGATCTGAATGCCACATCCGCCCGGGAGGAGACGCACAGTGCGCTTAGTAACTTGGAGCGGCCCCGC
GGCATCACCTTCGCCGCGGTGCCCAGTCCGCTGCCCAAGTGTCCGCTCTGCGGCGCCGACATCAGCAGCACCACCGGCACCACCGCAAACGCAACCGCAACCGCCAACGCCG
ACAGCACAATCGACACCACAGTAACCACTAGTAGTAAACGTAGTATTCACGAGCAAACTCCAGATCTTGGCCAGAGGCCCGCCTCCTCGTCCAGTTCCACTCGATTCTGGCA
CAAGCGAACGGCGGCGGTCACAGCATGTTGGCAGCAGAGCAAGAACCGGAAGCGCCGCTTCAAAACCGGATGCAGTCACTGTGGTGCAACCGGCGGTTCCGTGCGACCGGCC
AAAGGATGGAAGGCCGAACACAAGGCCGCCCGCACCCTGGGCATCATCATGGGCGTCTTTCTGCTCTGCTGGCTGCCCTTCTTTCTGTGGTATGTCATCACATCGCTCTGCG
GTCCGGCCTGCCCATGTCCCGATGTGCTCGTGGTGGTGTTATTCTGGATCGGTTACTTCAACTCCACGCTAAATCCGCTTATATACGCCTACTTTAATCGCGATTTTCGCGA
GGCATTCCGCAATACGCTGGAGTGTGTGCTGCCCTGTCTGGAGAAACGAAATCCGTACAACGCCTACTACGTCTAG
(SEQ ID NO: 188)

Start ATG: 281 (Reverse strand: CAT)

MSGVNVADLLATTMTLPITAAAGAATSQAAATSATNASHLQPATLTGHISTTAAAKTTTTPTSSLPITSQFVDASLTSLSLTATSSDASYSSPFSSYLSSDSTFELLSTVGP
NITANGSDIAVDNQAELEESWLDLSLLLLKGFIFSSIILAAVLGNALVIISVQRNRKLRVITNYFVVSLAMADMLVALCAMTFNASVELSGGKWMFGPFMCNVYNSLDVYFS
TASILHLCCISVDRYYAIVRPLEYPLNMTHKTVCFMLANVWILPALISFTPIFLGWYTTEEHLREISLHPDQCSFVVNKAYALISSSVSFWIPGIVMLVMYWRIFKEAIRQR
KALSRTSSNILLNSVHMGHTQQPTSLSYLHPSDCDLNATSAREETHSALSNLERPRGITFAAVPSPLPKCPLCGADISSTTGTTANATATANADSTIDTTVTTSSKRSIHEQ
TPDLGQRPASSSSSTRFWHKRTAAVTACWQQSKNRKRRFKTGCSHCGATGGSVRPAKGWKAEHKAARTLGIIMGVFLLCWLPFFLWYVITSLCGPACPCPDVLVVVLFWIGY
FNSTLNPLIYAYFNRDFREAFRNTLECVLPCLEKRNPYNAYYV*
(SEQ ID NO: 189)

Name: Beta Adrenergic/ Octopamine Receptor-like
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384518
GTTCCGTTACTTGGTCTTTGACTTCTTACCCCAGAATATGATTCTTGGCAGCAGGAGAAACCAGTACGACTTGTTTTGGTCGTATTATGAGTGAAATATTACCACTCATGTC
ATTTTGAACGTCTTTAAAGCTCACACCCTTATGGACCCATTATAAGAATATGGCGTGGCAAAATTTCCCATTTTTATCGAGGCTGGAAAAACTTTTCCACACATGCAATCAT
GTTTCGTTTCACTTATCATGGCTAAAACCTAGCAAATAGAAAAAGTTTTTTTTATCCACAAACATATCTGTATTGAGTGCCAGCAATCTGTATTGCTTTAGCTTAAGATGT
TTACTATTAATAGAGCCAAGTGGCAAATGACATACGGCTTACAGAGAAAATAATTTCCAATTCAATTACATTTACGTACATATATTCCTTCCTACAAGCAGACAGGATTATT
GATCTAAATATAAATAAATATATAATAATTATTTAATAATATAATATTAAAATGTAGATTAACAGAATATATTTTTTTTGTATGAAAGCCTCTAGGGTATTTGTGTGACTT
TTTTCTCTCTGCATAAGTGAAGATGTTTACTATTAATAGAGCCAAGAGGCAAGAGAAAAGTGGCAGCAATCAAGTGCTTTTATTTCGGGACTGCATGTTAAAGGCGATTTCC
TGATCGAAGACTTTGGATACCTGATGCCATCCGTCTGAATGTGACGTCAGCCTGGCTGGGGATTCAATTTAAATTCATCCTTGAGTCTAATGACTTTTAATGTTGTTTGCTT
CACAGGCGAGTGCGCAATTTTTGGGCCTGTCGTCGACTGCTTATCGGCTATGTGCCACCGCAGCCGCAAAGGATACGACAGACAATATTAATAACGCATTGATACGAGTGGG
CGGCCGATTTGCTTTTGTTTTTCTTTCCCAACTTGGCGGTGGAAAGGAGTGGTGTGGCGTGGCGTGGAGTGGATAGAGTCCACAATTAATCGACGACAGCTAGTATGTTGCA
AGGCGTCGCCATCACCATAGCCAACGACAGCAACGATGATGGCATCAATCAATCATTCATGGCTCATGTGTCGCCCAGTCCCAATCAGAGTCCCTCCATTGGAGTGGGCATC
GGTATCGCCTCATCCACAATGGCGAATCCCAGCGAGAGCCCCGAGATGTTGCTGCTGAAGAACGACAAATTCCTAACACATGTCGCCCATCTGCTGAACATAACGACCGAGA
ATCTTTCAAATCTCCTGGGTTCCACGAACGGCACAAATGCGAGCACAATGGCAGCGGATTCGCCGGTCGACGAGTCCCTGACCCTGCGAACAGCCCTCACCGTTTGTTATGC
ACTCATTTTTGTGGCCGGCGTGCTGGGTAATCTTATCACCTGCATTGTTATATCGCGAAATAACTTTATGCACACGGCCACCAATTTCTATTTGTTCAACCTGGCGGTTTCC
GACTTAATACTGTTGGTCTCAGGTGAGATGGCATTTTTGAGGGACACCATTGCGTATGCGTAATGCGGCAAATATCCTGCAGGCATCCCCCAGGAGCTGTACAACCTTTGGT
ATCCGGATATGTATCCCTTCACGGACGCCATGTGCATCATGGGCAGTGTGCTCTCGGAAATGGCCGCCAATGCAACGGTTCTCACCATCACAGCATTCACTGTGGAGCGATA
TATAGCAATCTGTCATCCATTCCGGTGAGTTTTTAACATTTAAACGTTATATATATTTTAAATATTTATTTACATTATTAAACAAATGTATAATTTCTTCAGTGTTATTGCC
TCGCTCTCATTAACAAAGTGGATCATGCATAATTAGGACAGAGAATATGAATAGATAAATTAAACTTGATAAGCGTCAGTCGGAAACATTATGATATGAAACAACAAAAGTA
CCTCACAAATTATGGGCACAATAATTATTCATAACTAAACAGTCTGGCTGTGGAATGATAAAACAAGTTAGGCTTGACAACCAAATTCCCATCGTTTTTAAATATACCAAAT
TCATTAGCTTTCATTTATAGACACTAATTAAAGTCAGGTCCATAAATGAACCAAGTTCATTACAGACTATGTAATTAAGTTCCAATAAATGCGTTGTAAAGATAGAAAGTTA
GTTAGTCAATTAATTGCTTAGTACCAACATATTTGCATTTTTGGAATGGTACAACAAAACTTTTCCTTATACAGGTATCACCAGTTCCTTCCAAATGTTCTGACCATTTTGC
GGTCAAAGAATACAACAGCTTCCTTAAGCGAAAGCGTCTGGTTTTCATTTACGATGTTGCATTTCCTTTAGTTCAGAAACTCTTGAATAGCTAGAAGAAATGAAACTTAAAA
ACCTTACATTGTTGTTTCAAACACAATTTCCGTTTCTTAGCATTTTTCACAAATTAATTGTTACGAATACAATCACTTGGAGTATGTCCTTAAGAATATGCAGACACAAAA
GGGTGGTTTTCTATAAGAAGGGATTCCAAGTCCGATTACAAAGTTAAGGCAAAAAACTTTCATGTCTCTGAAGTTGGGATTAAATATTGTTGTCGAAAAGAGTAAAAGGTTT
CTATGTTTGCCATTTAATAGTATTATAGTTATTTGAAACGAAAGGATATTCGGACATGTGCTTAGCATGAAAAAGGTCTGAAACCCATTGCAGTCACTTTACCAAAAGAAG
TTAAAAGTGATGTCTATTAAATTTGGTATGAAACTGTTTGAATTCCTTCCCATTTACACCCACATTCAAAGTAAATATTTATGTTGCACTCAATATCCCAAACCTTAAAGAG
AAAGTTAAAAAGCGTAAAAAGTGTAGGTGGTCACTGCGTGTACTTTACAACCGTAATATTTTCCATCTTATATCTGCCCACGTTTATCCGAAACATCGACATCATCTTGTCC
TTAATACTTAACATATAATGTAGCCTAATCTTTCCGTCATTTGTAGGCAGCACACCATGTCAAAATTGTCGCGAGCCATTAAATTTATATTTGCCATTTGGCTGGCGGCCTT
CCTTTTGGCCCTTCCGCAGGCCATGCAATTTTCGGTGGTCTACCAGAACGAAGGATACTCGTGCACGGTGGGTGGGACTGAAATGTGAAAACCAACTGACGCATCATCGAGT
GTCCTTTTCGCTGTTTATTCCCTTAGATGGAAAACGACTTTTACGCCCATGTGTTCGCCGTTTCGGGATTTATTTTCTTCGGTGGACCCATGACGGCGATTTGTGTACTGTA
CGTGCTGATCGGAGTGAAACTGAAGACGAGTCGACTGCTGCAATCGCTGCCGCGAAGGACTTTCGATGCGAATCGCGGCCTAAATGCCCAGGGACGAGTCATCAGAATGTTG
GGTAGGTGAGTTAAGTTGGCGGCTTGATTGATTAAAGACAGCAATATGCCCCACCGCCAACCAGAAACAACAAGAACATCCACGGCTGTCAAAACTTTTTCCGCCGAAAGTC
TTAAATTGAAAATTACTGTCAACGTATTTGCGTTGGCGAACATCAAACAGGGAATTCGTCATAGCCCCGCCCTCATTGCCTTCCGCCCACTTCATTTGCACTTGCGTCATGT
GGGCTGAACGTGCTTCTGTTTTTTTCCCTCTTTTCCCTCAACGGTGTCGTTAATTTTTCGCAAACATCGGTGCGTATGTGTAAAGCTCTTAAATGGCCATCTAGCCTATGAA
GCAAAGGTGAAGCCAGGTGTCACCTCAAAAGTAATCAAGCCGGGGCACTTAAATGAACTCAAGACTGCTGGGTGCACACAGCAAGCTTAAATATTTGTAAAAAGTGTTTCTT
GAATCTATAGACATATGTCCGTTAACGTGCCACGTTATAAGCTGTTAAATTTTAAAGGAATTCAGCACACATCTCGTGTATTATTCACATATTTTATTTATATAATTAGAAA
TTCTTATCAAACCATACAGTGAATAAATTTAAGTGTAAATCCAGGCCCAATTTTAGATCCCAATAAGCCGTTTTTTCTTACAGCCTACGAAACAACGTTTAATTTACTGTTG
CTTTTAACTGAAACACCGACCATTGATTTCGAAATTTACCATTGGCCCTGGCCTGCCAAATGGAAATTGATGGACCAATTCAGATTAAGCCCAGAGCCTTAACCATAAATTA
ATTTGGCCAGCAAGAGTGGAAAAAAAACACACCCGTTCTCAGGGCTTAGTTATTCATAGATAGTGGGATATTATTTATCGTTTGCTAAGAGGTAATACCTTGGCTTTTATCA
AAATGTGATATTTTACCAAAGGCCCGTTTGCGCAAACTATTCGATTTCGGGGGATATAAAAGATTAAGACCACTTTCAGTGTGAGAGAGCCCCTTTTCCGGTGACCTAGTTG

TTCAGGGGTCTTAATACGTTGCTTATCCTTACTTGACCGTATTTGCAGTAGCTGTAGCCGTCGCCTTCTTCCTCTGCTGGGCTCCATTCCACGCCCAGCGACTGATGGCAGT
TTACGGACTGAATCTGATTAATATAGGCATCAGTCGGGACGCCTTCAACGATTACTTCCGGATACTTGATTACACGTCCGGGGTGCTCTACTTTCTCTCCACCTGCATCAAT
CCGCTACTGTACAACATCATGAGCCACAAGTTCCGCGAGGCCTTCAAGGTAAGCCCATGGGATCAACCCTTCTCAGTTAACCACTTCATAATAGTAATCTATTATTAATTGG
GTCCTTCGCATGTTTGTGAGGACGGTCACAAATGCAGATGCTTTTAAAGTTAAATTTAAGCTTTAATTATTCAAAATTGTTTTTCGTAAATTAATGAATTGGTAATCTATGT
ATTTTTTACACTTTTAAGATCACTCTGACGCGACAGTTTGGTCTGGCCAGGAACCACCACCACCAGCAGAGTCAGCACCACCAGCACAACTACAGTGCCCTTCTCCGCCAGA
ATGGATCAATGCGTCTGCAGCCGGCCAGTTGCAGTGTAAACAACAACGCGCTGGAGCCATATGGATCCTACCGAGTGGTTCAGTTTCGTTGCCGGGATGCGAACCACCAATT
GTCCCTGCAGGACAGCATTCGCACCACCACCACAACCACGACGATAAACAGCAACAGCATGGCAGCCGGAAATGGAGTTGGTGGTGGTGCTGGTGGCGGTGGCGGTGGGCGG
CGTCTGCGGAAACAGGAGCTCTATGGGCCAGGTCCGGGTACAGCAGTTCCCCATCGAATGCTGCAGGCCCAGGTGTCTCAACTCTCGTCGCTGGGCGATGCCAATTCACTGC
TGGAGGCGGAGGTGGTGGACAGACACTACGCATCCGGGCGAGCCAAAAGAGCTCTCTTGGCCACCAAAAGTGGTGCCCTGCTGGTGACACCACCTCAAAGTGGTGACCCATC
GGAAGTCAGTCAGCCCGCCACTCGTCTCAAGCTGACTAGGGTGATAAGTAGAAGGGATGAGGTGGCCAACACAAGCACTCCGCCCTTTTGTGGGAGCCACAGTTTGCCGGAT
CCGGAGACTTGTCAATCGGCATCGGTGGCTGGTCGAAGCTCCCGTAAGTTTCCGTGGCGAAAACGAAGGCAAAAAACGGAGGATCCCAGCAGCGAGGGCTTGACCTATGGCT
CACCCAAATCCCAGTGATTTCAGCTGGCTGGTAGCTAGAGTTATACTTAGTTAGGCATTGTTTAGTTAGCGCAATAAACTTTTCTATCTGTAAATAAGCCAGAAAAAGTCCT
GCGATTTGAATGCAAATTTCGGGTGACCATAAAAAATTTAACGAATGGGCATGGTCCAGACAATTTATGAGAGTCATAAAGCCGGGATAATGAGGCTACGGCGTATCGTGAC
GAATACACGCCCCACATCATTTTGCAATGGCAACCGCATAGCACTCATAAATTAAGTGTTAAGCTACCAGCTAAATAATGTGAATAAATAAAAATTGTTCTCAATATTACAG
CAAAATCCTTTTAAATGCGTTTTCCTTTCACAATGCATGCCAGTTACATTCAAGTTCGCCTGCATTCCAAGTTGTCCAGGATCAAAGGGAAAATCAAAAAAAACAGTTTACA
TCTCAAGTTTTGCAGGTGCAGCGTTTCCAAGCAATCTGTTCCTACCTTTTGCGACATTTTTCCTGCAGATTGCCTGGCATTCTGTTGATTTCCCACACAAAAACTAAAAGCG
CTGAAAAAAATAAAAGAAGCATAAAATAAAATAAAGTGCCAATAAATTAGCAGTCTTAATATTTTCAGAGCCCCTCAGAAGCAAATCTAGGAAGGAAAACATGACATAAATT
TTAAATGGATTGCTCATCAGGATATCGTGTTGCAGTAGATATATATATTTTCGCAGTGTAAGTCACGACTCCACTTGCACCTGGCAACAAAGTTTCAGTTTGCACCTGGCCC
ATTGCCGGAGGTCCATTAGTCAAGAGGGCTTTGTAGGCTAGGAAGTCCGGATTGAAACGAATGTGCAGAGAGTGGCACGAATGGGCTTATCCTGCGGTGAATTGGCACGTGG
AGTGCATGGAGAGTGGTGTCGGTGCCACAAAATAAATCAACTTAAAAGTAGCATACAAATTTAAGGACCCGACTGATGTTATGTCCAAAGGGAATTATTAATGGTTCTTTAA
AGGCAGGCA
(SEQ ID NO: 190)

Exon: 1001..1478
Exon: 1539..1704
Exon: 2959..3091
Exon: 3163..3361
Exon: 4417..4640
Exon: 4835..5617
Start ATG: 1001

Transcript No. : CT25324
ATGTTGCAAGGCGTCGCCATCACCATAGCCAACGACAGCAACGATGATGGCATCAATCAATCATTCATGGCTCATGTGTCGCCCAGTCCCAATCAGAGTCCCTCCATTGGAG
TGGGCATCGGTATCGCCTCATCCACAATGGCGAATCCCAGCGAGAGCCCCGAGATGTTGCTGCTGAAGAACGACAAATTCCTAACACATGTCGCCCATCTGCTGAACATAAC
GACCGAGAATCTTTCAAATCTCCTGGGTTCCACGAACGGCACAAATGCGAGCACAATGGCAGCGGATTCGCCGGTCGACGAGTCCCTGACCCTGCGAACAGCCCTCACCGTT
TGTTATGCACTCATTTTTGTGGCCGGCGTGCTGGGTAATCTTATCACCTGCATTGTTATATCGCGAAATAACTTTATGCACACGGCCACCAATTTCTATTTGTTCAACCTGG
CGGTTTCCGACTTAATACTGTTGGTCTCAGGCATCCCCCAGGAGCTGTACAACCTTTGGTATCCGGATATGTATCCCTTCACGGACGCCATGTGCATCATGGGCAGTGTGCT
CTCGGAAATGGCCGCCAATGCAACGGTTCTCACCATCACAGCATTCACTGTGGAGCGATATATAGCAATCTGTCATCCATTCCGGCAGCACACCATGTCAAAATTGTCGCGA
GCCATTAAATTTATATTTGCCATTTGGCTGGCGGCCTTCCTTTTGGCCCTTCCGCAGGCCATGCAATTTTCGGTGGTCTACCAGAACGAAGGATACTCGTGCACGATGGAAA
ACGACTTTTACGCCCATGTGTTCGCCGTTTCGGGATTTATTTTCTTCGGTGGACCCATGACGGCGATTTGTGTACTGTACGTGCTGATCGGAGTGAAACTGAAGAGGAGTCG
ACTGCTGCAATCGCTGCCGCGAAGGACTTTCGATGCGAATCGCGGCCTAAATGCCCAGGGACGAGTCATCAGAATGTTGGTAGCTGTAGCCGTCGCCTTCTTCCTCTGCTGG
GCTCCATTCCACGCCCAGCGACTGATGGCAGTTTACGGACTGAATCTGATTAATATAGGCATCAGTCGGGACGCCTTCAACGATTACTTCCGGATACTTGATTACACGTCCG
GGGTGCTCTACTTTCTCTCCACCTGCATCAATCCGCTACTGTACAACATCATGAGCCACAAGTTCCGCGAGGCCTTCAAGATCACTCTGACGCGACAGTTTGGTCTGGCCAG
GAACCACCACCACCAGCAGAGTCAGCACCACCAGCACAACTACAGTGCCCTTCTCCGCCAGAATGGATCAATGCGTCTGCAGCCGGCCAGTTGCAGTGTAAACAACAACGCG
CTGGAGCCATATGGATCCTACCGAGTGGTTCAGTTTCGTTGCCGGGATGCGAACCACCAATTGTCCCTGCAGGACAGCATTCGCACCACCACCACAACCACGACGATAAACA
GCAACAGCATGGCAGCCGGAAATGGAGTTGGTGGTGGTGCTGGTGGCGGTGGCGGTGGGCGGCGTCTGCGGAAACAGGAGCTCTATGGGCCAGGTCCGGGTACAGCAGTTCC
CCATCGAATGCTGCAGGCCCAGGTGTCTCAACTCTCGTCGCTGGGCGATGCCAATTCACTGCTGGAGGCGGAGGTGGTGGACAGACACTACGCATCCGGGCGAGCCAAAAGA
GCTCTCTTGGCCACCAAAAGTGGTGCCCTGCTGGTGACACCACCTCAAAGTGGTGACCCATCGGAAGTCAGTCAGCCCGCCACTCGTCTCAAGCTGACTAGGGTGATAAGTA
GAAGGGATGAGGTGGCCAACACAAGCACTCCGCCCTTTTGTGGGAGCCACAGTTTGCCGGATCCGGAGACTTGTCAATCGGCATCGGTGGCTGGTCGAAGCTCCCGTAAGTT
TCCGTGGCGAAAACGAAGGCAAAAAACGGAGGATCCCAGCAGCGAGGGCTTGACCTATGGCTCACCCAAATCCCAGTGA
(SEQ ID NO: 191)

Start ATG: 1

MLQGVAITIANDSNDDGINQSFMAHVSPSPNQSPSIGVGIGIASSTMANPSESPEMLLLKNDKFLTHVAHLLNITTENLSNLLGSTNGTNASTMAADSPVDESLTLRTALTV
CYALIFVAGVLGNLITCIVISRNNFMHTATNFYLFNLAVSDLILLVSGIPQELYNLWYPDMYPFTDAMCIMGSVLSEMAANATVLTITAFTVERYIAICHPFRQHTMSKLSR
AIKFIFAIWLAAFLLALPQAMQFSVVYQNEGYSCTMENDFYAHVFAVSGFIFFGGPMTAICVLYVLIGVKLKRSRLLQSLPRRTFDANRGLNAQGRVIRMLVAVAVAFFLCW
APFHAQRLMAVYGLNLINIGISRDAFNDYFRILDYTSGVLYFLSTCINPLLYNIMSHKFREAFKITLTRQFGLARNHHHQQSQHHQHNYSALLRQNGSMRLQPASCSVNNNA
LEPYGSYRVVQFRCRDANHQLSLQDSIRTTTTTTINSNSMAAGNGVGGGAGGGGGRRLRKQELYGPGPGTAVPHRMLQAQVSQLSSLGDANSLLEAEVVDRHYASGRAKR
ALLATKSGALLVTPPQSGDPSEVSQPATRLKLTRVISRRDEVANTSTPPFCGSHSLPDPETCQSASVAGRSSRKFPWRKRRQKTEDPSSEGLTYGSPKSQ*
(SEQ ID NO: 192)

Name: GROWTH HORMONE SECRETAGOGUE RECEPTOR-like 2
Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384518
TTGGGCAAATTAAGGAAACTTTTTGTGGTTTCAGTTACTCGTTTAAGCCACATTTTCCCAACATATGCTGCACTTTATGAATTAAAAGTCATTTATTTTAGGCTGCACACAC
GTTATGTTTGCCTGACAAATTGCATTTGCTGAAATTAAAACAAAATGCACTGGGACACTGCGTTATGGCATGCTGCTAATTAGTCACCAACAGCGGCTTTCAGCGAATGTTG
GCCAATCGCGAGAATGGCTTTCATCCACTGTCTAGACTCCAGACCAAAAAGCGAATTTGTGCGCCAGATAAACATCTTAATGGCCAGGGATTATGATGCAGTGCCCGCTTTT
TACTGACAGTAATTAGCCAGACGCCCATTGAACCCTCTTCTTGATAATCCCCACTTGATCCACACTCTTTCAAAAAAAAAAAAATGTATTCATTAAATGGGAAAAAAATATTT
ACTGCAAGAGAAATGTTATTAATGGATATTGAAAAACTGACATTAGATAACAGTTTCTATTGGACTGCTAACAAAGTCGGTCAGCTGTTTAAAAAACTGAGATCGCTTTTAT
ATTTAAATTGTATAATAAAACCATTAGGGTATCAACCTTAATATGTATTTATATATCACTGTACTCCAAATACTATTTACGCTTAACTATGTAGATTCATTTTATTTTTATC
AAAAGTCCATTCTTTTTTTATTAAACTCCACTTTATAAATGAAACAACGGTCGAATTTTTGAACTGAGTACATAACTACATAGTTTATTCGATTACGTGCACTATAACTATA
ATTCGAACTATAGCTTTGGGCTATATTACGGTTATTGGTAAGGACAACCATATAATTGACCCCAAACGCGGAATATTTGTCTAAATTAAAACACAAACAAGCTGGAGTGGCG
CTTGCTGCTCGTATCACAGTTGTGTGTGTGTGTAATGTACGTAGGTGGCACTGGAATGGACTACCTCGCTTATCGAGTACAGATAGAGGATATGCTGACCCCCTTTAAAAGG

CGGCCCGCTCTTCAGTGTCGCCTCCATTCACCGTCGCCCCCTTTTCGTGCGGTTTGTGGCTGCCTTTGCGTCTGGCACGCCAATTGAATTTGGTCCTGAGACCCTTGAACGG
GCTTGAGCTCTTGAGCACTTTCACTTTCGCCTTGCGCACTTTGCCACTCGACTCGTCACCGCTCACGCCCCCCGACCCCGCGGCCCCTGTTGTGCTGCCTCCTGCCACGCCC
CCCGTCACACCTCCTGGTGGGCGGGTGATGCGCGTAAGTTTAAGGCGTGACCGCGAGATGGCCAAGTCATCTGTGAGCTCGTCAATGGAGCACATCGTGGGCTGGCGCTGCA
GTGGCTCCCCGGGAAACTGCAAATCCTCCTCCAAGAGGGAATGGGCATCGCCCACCGATGACAGCTGCGATATTTGCGTCTGGAGCATGCGACGCGGATTATCGAAGAGGTC
TTGCCGGCTGCGATTCTGACCCTGCATCTGGACACTGTCCAAGGACACGCGGTTCAGGCGCACCGACTGACCCGTTGTCGTTCCATTGCCACTGCCATTGGCAGAGCCATTT
AGCCCAGTGGTGGTCGTCGCCATGGAAGTCATTGTGGTGCGAACGCTGTCCTGCAGGAGGAAATCAAGGAATGGAGGGATAATCAGAAAGGGAAAGTGCAGTCGCCAGACAC
GAAAATAAGTAACGCAATAAACATCTTATAGTGGTTTTAAAAAGTCTCGTGATTCAAATTTCATTTTGAGAAACAGGACATACACTATATTAATCCTTTTCATTTTCCTCTT
AAATAATGCCACATAATAGAATTTAAGTGCATATATCGAATGATAATTATGTAATTTTTCTCAGTGCACTCACCGTTGTGTGCAGCCGCAGTGAACCCGTTTGATTACGTCG
AAGAGCGCTGTAAGTATGAGGAAGCCCCCTTCCCTGATTTTTGCCACCAAGTCCAAAGTGTCTAGCCAAAGTTACCTAAAATGGGATTAAATACGTTTTTAATTCTAGCTT
GAAGTTTAAAATCACTAAATTGACCGTTAAAGAATCGATTAATAAACTTGTAACCTGAAGCTTATTTTCAATTTCCTAACGCAAGCTATAGAAATGAATATCCGGCAGCGAA
ACTCACCTTAAAAGCCTCACGAAACTTGTGGCTCATGATGTTGTAGAGCAGCGGGTTGATGCAAGTTGAGAGGAAGTAGAGAACACCGGACGTATAGTCGAGGATGCTGAAC
ACGTCGTTGAACCACTGGGACTCAATGCCCGAGGTGGATCCATAGACCGCCATCAGCCGCTGGGCGTGAAAGGGGGCCCAGCAGATGAAGAAGGCCACGGCCACCGCCACTG
GGAAAGTTGGATAAGGGCGCCTGATGCAACCATGAGTTATTTCGGATATTTCATATCCTGCTTACCCAGCATCCGGATGACTCGCGTTTGGGCGCTTATCCCCCGGTTTACA
TCGTAACATCGCCTGGGAAGCGCCTGCAGGAGTCGGCTCCGTTTCAACTTCACCCCGATGAGGACATAGAGCACGCAGATGGCCGTCATGGGTCCGCCAAAGAACAGGAAGC
CCGACACAGCAAACACATGGGCAAAAAAGTCGTTTTTCATCTACAGCAGGCATTTCACAAAAATAATATATATGTGTATTCGCATGCGATACGATATGGTTTTGGCTTACAG
CCACAGCCGGAAGGCAATGGCTGTAAATTTGATTGCCTAAACTGAGTCGAACCTCGTAAGGGATTGTCGCTGTTTTTATCATGCTCTTAAATTTTTTGTAATTCACAAAGGC
TCTCTATTCGAACTTTCTATACGGAAAACACATAATGAATCAAAGGATCAAAGGACCGGTTCTACGAATCAATTAACCAGGTAACTCATAACTGGGGGACATACGATTAAGG
GTCTCGCATCAATCATACGCACCGTGCACGATGTTCCCATGCCCTGCATCACCACGGAGAACTGAATGGCTTGGGGCAGGGCCAGCAAAAGGGCAGCTATCCAGATGGCAAA
TATGAACTTTACGGCCCGTGACAACTTGGACATCGTGTGCTGCCTGTATCGAAAGCAGAACCAATTAAGCAAGTTTGGGGCAAAATGGAAATGGGCAGTGGAAGGAATGAAG
GTTCCCTTACCTGAACGGATGACAAATGGCAATATATCGTTCGACTGTGAACGCGGTAATGGTTAGAACTGTCGCATTGGCCGCCGTTTCCGAGAGAACGCTCTCCAATATG
CAGATGCTGTCACTGAAAGGATAATTATCCGGGTGCCAGAGGTTATAGAGGTCCTGCGGCATTCCTAAGAACATAAATAGCAAAATAGCAAAATATATAGTGTGATTATAAA
TATTTAATAATAGCTGAAATTATAGCCTATAGCCTCAAAATGCAAATCGATTTGAGCATCTGCCGAATATCTCACTCTACATTTTCTTGTTTATTTACAAATTTTAATTTAA
TATGCCTCAGTTCCATTTATTGTTTGCACTACTGAACTGAATTAAATTATAATTCAAAAACCATGAACATTTTAGTTTTGCACTCACCTGAGCATAACAAAATCATGTCGGA
TATAGCGAGGTTAAACAGATAAAAGTTGGTGGCCGTGTGCATAAAGTTGTTCCGCGAAATGACAATGCATGTGATGAGGTTGCCCAAAACGCCGGCGATAAATATGAGGGCG
TAGCCCACGCTTAGGGTGGCCAGCAGTGAAAGTGGTGTCATCGGGGTCACCGTGGGCAAAAGCTCCTCCGGCTCCAGGCCCTGGAGAAGGCTGGTCAGGTTGTCCGCGGAGA
TGTTCAGCACTTGTGTCAGATTTAAGAGAAATTTCTCATTGTGAAACAGTTGTAGATCGGTTGCCAGGACACCGGAACTGTTAGTGGGCAGCATTTTGACTGCCATAGGGTG
TGTGTCTTTGTGTGTGTGTATACGCACTATTATTTTCGAATTTCTTTGGAAATTTCAGACTTCTGCGTATGTATTCTCAATCGAAGCTTTCGCTTGCAAATAATGTTTTCAT
AACGCCCCTTACGGGGGCTCATCAATTTCGACTTAGGCCAGAGCCACTCCACTTGCAATCTGTTGTCATGAGTGTGGGATTAAATTATGTGTACGGATGTGTGTGATAAGGT
GGACCCTATGTGCTATTTAAATAAAACTATGTTTTTCAAAATTTGTTAAAATTTAAATAAAATAATACAAGACAAATAACGCATTTTAACGAATTTAGTAATAAAACAAATT
CAAACATTAAGGCTTTTTACTCCAGATTTCTAAATTACCATAAATTATACATATACACAATTTGGTTGGTAAGCTGTGCGCTTTGTTTAGCTTGCTAAGGTCCACCTTAGAG
AGCCCGTCTTAAAGCGTGTTTGTGTCCTGTGTGTTTACCTGGCATCCTGGGAAACTGGCAGGCGGCAGTCAGGCGTGTCCTAGGCGAAAGACCAGGATCAGGGCGAGTGGAA
GCTGAAAATTAATTTGAATTGCTAAGTGGCATTTAGTCAAAATATTTGTGTACACTGGATGAATTTAAGGATGTTGCTACAAGGCTGAGAAGGACAGGCCTGCCGATCCGTC
TCCCACCCAACACCAGGTGGCAAACGAATGAAACGAGCCAGAAGCTCAAGGTCAGGTGGTGGATAAGACGGATAAGGAGACGACGATGGCAAATGGACACCGGGTGACTTAA
TAACATTAACATAGTTTCATTCCATTCAGATGATGAGTGTTTTCGGCTGCTGCTGTTCTGCTCTGTTCTGTTCACATTCTGGCTATACTTTTCGTTATTGTTGCTGCTGACC
GCTGTTGACCGTTTTTGGACGTTTTCATTATCATTATCATAAAAAAGTTTCTACAATATATGCACAACCAAAAGAATGCTGTCGAA
(SEQ ID NO: 193)

Exon: 3902..3560
Exon: 3312..3147
Exon: 3067..2935
Exon: 2616..2418
Exon: 2349..2135
Exon: 1979..1866
Exon: 1618..1001
Start ATG: 3902 (Reverse strand: CAT)

Transcript No. : CT25350
ATGCTGCCCACTAACAGTTCCGGTGTCCTGGCAACCGATCTACAACTGTTTCACAATGAGAAATTTCTCTTAAATCTGACACAAGTGCTGAACATCTCCGCGGACAACCTGA
CCAGCCTTCTCCAGGGCCTGGAGCCGGAGGAGCTTTTGCCCACGGTGACCCCGATGACACCACTTTCACTGCTGGCCACCCTAAGCGTGGGCTACGCCCTCATATTTATCGC
CGGCGTTTTGGGCAACCTCATCACATGCATTGTCATTTCGCGGAACAACTTTATGCACACGGCCACCAACTTTTATCTGTTTAACCTCGCTATATCCGACATGATTTTGTTA
TGCTCAGGAATGCCGCAGGACCTCTATAACCTCTGGCACCCGGATAATTATCCTTTCAGTGACAGCATCTGCATATTGGAGAGCGTTCTCTCGGAAACGGCGGCCAATGCGA
CAGTTCTAACCATTACCGCGTTCACAGTCGAACGATATATTGCCATTTGTCATCCGTTCAGGCAGCACACGATGTCCAAGTTGTCACGGGCCGTAAAGTTCATATTTGCCAT
CTGGATAGCTGCCCTTTTGCTGGCCCTGCCCCAAGCCATTCAGTTCTCCGTGGTGATGCAGGGCATGGGAACATCGTGCACGATGAAAAACGACTTTTTTGCCCATGTGTTT
GCTGTGTCGGGCTTCCTGTTCTTTGGCGGACCCATGACGGCCATCTGCGTGCTCTATGTCCTCATCGGGGTGAAGTTGAAACGGAGCCGACTCCTGCAGGCGCTTCCCAGGC
GATGTTACGATGTAAACCGGGGGATAAGCGCCCAAACGCGAGTCATCCGGATGCTGGTGGCGGTGGCCGTGGCCTTCTTCATCTGCTGGGCCCCCTTTCACGCCCAGCGGCT
GATGGCGGTCTATGGATCCACCTCGGGCATTGAGTCCCAGTGGTTCAACGACGTGTTCAGCATCCTCGACTATACGTCCGGTGTTCTCTACTTCCTCTCAACTTGCATCAAC
CCGCTGCTCTACAACATCATGAGCCACAAGTTTCGTGAGGCTTTTAAGGTAACTTTGGCTAGACACTTTGGACTTGGTGGCAAAAATCAGGGAAGGGGGCTTCCTCATACTT
ACAGCGCTCTTCGACGTAATCAAACGGGTTCACTGCGGCTGCACACAACGGACAGCGTTCGCACCACAATGACTTCCATGGCGACGACCACCACTGGGCTAAATGGCTCTGC
CAATGGCAGTGGCAATGGAACGACAACGGGTCAGTCGGTGCGCCTGAACCGCGTGTCCTTGGACAGTGTCCAGATGCAGGGTCAGAATCGCAGCCGGCAAGACCTCTTCGAT
AATCCGCGTCGCATGCTCCAGACGCAAATATCGCAGCTGTCATCGGTGGGCGATGCCCATTCCCTCTTGGAGGAGGATTTGCAGTTTCCCGGGGAGCCACTGCAGCGCCAGC
CCACGATGTGCTCCATTGACGAGCTCACAGATGACTTGGCCATCTCGCGGTCACGCCTTAAACTTACGCGCATCACCCGCCCACCAGGAGGTGTGACGGGGGGCGTGGCAGG
AGGCAGCACAACAGGGGCCGCGGGGTCGGGGGGCGTGAGCGGTGACGAGTCGAGTGGCAAAGTGCGCAAGGCGAAAGTGAAAGTGCTCAAGAGCTCAAGCCCGTTCAAGGGT
CTCAGGACCAAATTCAATTGGCGTGCCAGACGCAAAGGCAGCCACAAACCGCACGAAAAGGGGGCGACGGTGAATGGAGGCGACACTGAAGAGCGGGCCGCCTTTTAA
(SEQ ID NO: 194)

Start ATG: 1 (Reverse strand: CAT)

MLPTNSSGVLATDLQLFHNEKFLLNLTQVLNISADNLTSLLQGLEPEELLPTVTPMTPLSLLATLSVGYALIFIAGVLGNLITCIVISRNNFMHTATNFYLFNLAISDMILL
CSGMPQDLYNLWHPDNYPFSDSICILESVLSETAANATVLTITAFTVERYIAICHPFRQHTMSKLSRAVKFIFAIWIAALLLALPQAIQFSVVMQGMGTSCTMKNDFFAHVF
AVSGFLFFGGPMTAICVLYVLIGVKLKRSRLLQALPRRCYDVNRGISAQTRVIRMLVAVAVAFFICWAPFHAQRLMAVYGSTSGIESQWFNDVFSILDYTSGVLYFLSTCIN
PLLYNIMSHKFREAFKVTLARHFGLGGKNQGRGLPHTYSALRRNQTGSLRLHTTDSVRTTMTSMATTTTGLNGSANGSGNGTTTGQSVRLNRVSLDSVQMQGQNRSRQDLFD
NPRRMLQTQISQLSSVGDAHSLLEEDLQFPGEPLQRQPTMCSIDELTDDLAISRSRLKLTRITRPPGGVTGGVAGGSTTGAAGSGGVSGDESSGKVRKAKVKVLKSSSPFKG
LRTKFNWRARRKGSHKPHEKGATVNGGDTEERAAF*
(SEQ ID NO: 195)

Name: GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1/like

Classification: G_protein_linked_receptor

Celera Sequence No. : 142000013384638

TAATTGCTGCCGGGTAATTGCAACTGAGCTCAGGTCAGCGAGTTTGCTAACCATGGCGTCATTATACTGTTGCACTATAAAATCATTCGACCCAAGTCCGATTAACTTAAGT
GGGTCTTAGACCAAAAGATGTGTATTTCGAATACTTCTCAAACGTTATTTCCTTAAGTCAAATCATTTCAAGAACTGGGAGAGCAAATAAATATGCGATTTCCTCGACTTAT
AGTTACTTGTGTGAAATTTGCGCGCTTGTGCAAATTCGAATGTTTGCAACTTAATTGGACTGCATAATAGATTGGGCAAAAGGATTTTATAACAATTCCAGTGTCCAGCCCA
TCGTACTGATGTGGCTAATTGTTTTGCTGACTTGCACCTAGGCAAAGCAATTTCCCTTGAAAACAATCAAATTCAAAACAATCAAGTAAACAAGACAGCTGAGTGAAAGACT
TTGGCTTCATGGTGAACATATTTGTCGGTTGAAGGTTGAGTTTTTGTTTAAGATGATTGCACATGGCACACAGGTTGATTTTAATCTAAATTTAATTTGAGCATTTGAGCA
TTTTTCTGCTGAACCGGGAAAACCCACATTTGCTTACCTAGAAAAACCGGTTTTCAGAGAAAGTTGCGAGTAATACATATTTAATACGATTTCAAGTCCTGTGTGTACAACG
TTTTATCTAAGAAAACCAGAAAACCAGAAGCTTTTTAGCTTAGACAGCTTTAATATTTAATTACAGTGTTCACAAAAAAAAAAAAACGAAATTCCCCTTTTCTTTGCTTCAAA
TTTAATTAGACATGTATCTATTGTTTTTTTTTTTCCACTTGCAGCTTATTTATCCAACCTAAGCCAACTAAATGATGCAATGCTGGTGACATATAAAACGAAAGTGGAAACG
GAGTTCATTCCTTAAGTAACCAATTTAAGATGTTGTGCTCCCAAGATCCCTAAACCACCGTCCAAACCCTTCGAATAACTTATTTAAGCCAAGATATAGACAAAATGGAGGA
CGAGTGGGGCTCCTTTGATCGTCTGCCCAGTGTTCCGAGTGCCTCGATGGATTTGGAGACGGAAAACGAGGTGGTCAGCAATTGGTCCACACTGGCCAACTTCACGCGACTT
GTGGCTGGTGCCGCTCCTGAAATCGTCAACTATACGCTCAACATGATCGACGTGGGTGTGGGCATGGCCACGGATATATCCAATTTGAGCGTGAGCACCACGCCCCTGCCCG
CCTACGCAATCTCCAATAGCTCCTCACTGGCGCACACCAATAGTCGCCATGAAGCACCACCGATGGCGGAACAGGTTCCGGAGCACGTGATGGATCACGCACCTCAACTATC
CCGATCCGGGTTGCTGAAAGTGTATGTCCTGGCGGTAATGGCACTGTTCTCTCTGCTGGGCAACCTGCTGACCATCTGGAATATCTACAAAACCCGCATCTCAAGAAGAAAC
TCACGGCACACGTGGAGTGCTATCTACTCACTGATGTTCCATCTGTCCATCGCCGATGTCCTGGTCACCTGGTTCTGCATTATCGGGGAGGCCGCGTGGTGCTACACCGTCC
AGTGGCTGGCCAATGAGCTCACCTGCAAGCTGGTGAAGCTCTTCCAGATGTTTAGCCTCTACCTGAGCACCTATGTCCTGGTCCTCATCGGAGTGGACCGCTGGATAGCGGT
CAAGTATCCGATGAAGTCGCTCAACATGGCCAAGAGGTGTCATAGGCTACTTGGCGGTACTTACATCCTGTCGCTGGTGCTCAGCTTGCCACAGGTAAGTTTTGATACACTG
AAAATAAGACCTTTAAAGTTCTTATAAAATCATTTGCCTAGAACAAACAGATAAGATAGTTACACAAATAAGCAAGATTGTAATAATATAAAAGTCATTTTTGGCCTATAAA
TGGTCTAAAAATATTTGTGGGATATGTTAGATTCAATCTTTAGATAATTTAAATGAAAAAAAGTTAAAGGTGCCCTTTGATTAGGAAATCAGTAAAAGCTTTAAGATCTCCT
CTTAATTTATGTACCCTTTCGGAAAATTTCTGTTCGACGTAAATGCAAAAAGCCTTGAGGGTACCCGTAAAAACATTATTCTATTTAAAAAGATTTAGCCAAAGGGCAAATT
TAAAGCGTCTGCATCTGAAAAAATGGCCCTTCCAGTGAAGTGCGAAGAAGACACGCTTATGAAAATAATCCAGTCACTTTTGGTGGGGCGTTCCATTACGCTAAGTTAAAGC
TGCGACTTAGCCGTGGTAATATTCACCTGCCACTGCCCCTGCCTGTGCCCCCTACTCCTGGCACTCCATTTGATCCTTACGCTCGCGCTGATGAAGTTGAAATAATATGTTT
ACTTTATAACACGTATACGCCCCGTGGCACAGTGCAAATAGTCTTGAGTAGGCGCCAGAATGTGTGCAGACATTAAAAATGATTGCCTTTCGCACAATAAGTGTACACTAAA
TGTGGCTTTAGGCCGCTGCCACGCCCACACACACGCGTGCTGATGGGCGTAGGAATGGGCATAGGGTTGGGCGTATGGGGACACCCAATCGAATTCCGGTCGTCCGACCGAC
AAAGGGTTTTTTTCTTCTGGTTTTTCCTCTCGGCTTTTCGCTTGGGGACGGGCGACCAAATAAAAAAAAAAAAAAGGTCCGAACAAGTGTGTCTTCTGCCTGACGACAGACAT
TATTCATATTCCCGATTTCGGTTCTTTGTGCGCAGATAGGCACACTGGGTGATATATTCCTTTGCCCTGGCAGTCAGCTGGCGGATAAGCCTGAATGGGACTGCATTAATGG
GCGTCACTCGCTTTGTTCTACAACTCTAGAACGACTCGGCTTCAGATTTCCGGCACCCTGATTGTATAATATATGCCGCGTGTCTCCACATCCCGATTTCATACTCGATTCG
CCTGGCCTTTCCTTGATTTATGGCCCCGGGTGAGGCACTTCACTTCTCGAATTCAGTTAATCTTTAGCTGGAGTGCTACGAGCTCCTGCTCGTGTTTTTCCTTAAAGTGGAA
ATTCGTGGACACTTCAGTTTAGGGTGTCACTAAGGTAGGTAATTAGCACGAAATCTCACTTCAAAGTTGCCAAGGGTTAGTAATATAGTTAGATTGAAAAGTGACTGCAAAC
TTTCCACTGCTCTTTATGTGACCCGGGTCATAAATATCGTTCGCAGAATTTGGTTAGCCATAAGTGGGGACAAGTGGGTAGAAAGGTCAAACCTTCTGTATAAAAAGATTTG
TCACCTGAGAGATGCCAAAAATAGATGAAAAAGAACCAAAGGGAAATTTCCTCCTAAGTGGATGGTAAGTGGATTTCCCGCACGCAAGTTATGCAACTAGCACACTTTCCCC
ACTGCTCGCCCGCAAATGTATGCTAAAATGTAGCTCATTTAATTAGGGAGCCCAGCAACCCGACGCCATTTCCTTTTATCTGAATAATGTATTGCGTGGAACAATAATTACA
GACGTACCGTACAAAAATCGAGGCGAGATTATATTTGCTCACTTTGGGGCAAAATGTGTGTGTGGTTTTTCTGGGTGCGAAGGCATGCGCGAGTCAGTTTGTTAAATCATCT
GACTTGTTTAAATGCCGCAAAAAATCTACTGAGCTTAATAACATTTTGATGTTTTTTTCCAATCGATGTCAAGCGGAAAATTTGGCCGTTTTGTTGCTGGGAATTTGGGCCG
ACTTTTGGTCCAGCCGACCCATAATGGAAAGGACATAATGTATTGGTTTTGGATTTTTAAAATAGAACCAGACTTTGATGAACTGTCCAGCTTGGGTGTAAATATTTGATGA
CGAACGTCACAAATTACGATGATATGACAGAAGTCATTTATGGATGGCCACTGACACCCAATTAAAAGAAACTTTGCAATCATTAAGCCAATTTTTATGGTCCACTTTTCGA
CATCTCGCTGAGATCCCGGGAAACTTTCTTCATTTCGCTTCTACTCGAATATGCGACATAAAATTTATGCCGCTTACAGACGTGACGGTCGAAACGTTCCCGCCGGAAACGT
AGGAATCATCTCAAACAATTTTCCGGATAACTTTTTCCATGTCACACGTTCATGGTGCAGGCAACCAAACAGCCAACCAACCAACCAAGTGGCACATGTATTGGTAAGGAGC
CACATCGAATCCAATATGAGACGCGCCGCTGAACTCGTCGGTTTCAAATAAAGGAGCGAAATGAAAATTCACTTCGACTGCCAATAAACAACAAATGTTTTCAACAGCAGCA
TTTATTTGGCTCCACTTCTATCGACCCCAACTCCTCCGCGTCCTTCGAACCCAAGGAAAAAAGAAAAAAAGTTTAGTACCACACGCACACAACCACACAGAAGAATCCTGCC
ACAATAACTCGAAGTGTGAGCCCAAAAATGTGGCACAATGAAAACAGGCCGCCATAATACAGTCAAACACAAAGTTATCCACAGGTCCTCCGCAGAGTATTATTGCGAATTG
AATTTCACCCACATATCATCAAATTCAGCAGCCTTAGGTGGCATAATTAAAATTTCGTAGAATTGTTTACAAATTGGAGGGAACTAGTCAATATTACAAGAGTCCCTTGGCG
ATTTCACCTTCAGTGTGCGCGGGCTTTTGATAATCGATAGGGGTCCTTAAATGCTTAGCTTTAAGTGAATTAAAATTTGAATTTAGTTTCAGAATATTTACAGCAGAATTTG
AATCAATAATTGCAATTTATATGTGCAAGCATTGGCGAATAAAGTAAATAACCGGATTTAATTATTAGGTGACAAAATCGTAACATTTAAATATATTTACCTGAATAACAAT
TGGAAATAAGTGCAAAAACAGCAAAAAGTATAATAAAGTTTAGATAAAAAATGAAATCCATTTAATATTTTAATTTGGGCTCATCTGGATTCATATTTTAGGTGACAAAATT
AATTATTATAATTATTAAAACAAATGAGCAGGTAACAAGTCAGTAAGCTCTAACAATCCTTGTTAAATCTAACTTAAACGTAGCTGAGTGTCCAATTTAAATGGTCAATTTA
ATCGTGCCTTTAAGCCAAACAAATATGTATATTTTAACAATCTGATAGTTTTACAAACGTCTTAAACGATTGCCAAGCACCTGACCGATGTGTGAAAACTACGAACTTGAGC
AATACAATTTGGTTAGAACGGAAGTTCTTGAAGTTGTTCCCTCACAATGAATCAATCTTGTTTTATTTTACTGTCGCAAGTTTTTGTCTAAAGTTTTTCGTTGCAACTTGTC
TAACTTTGCCGCAAATAGTGCTGAAGGCTTGCTTGTGTGTGGCATGGTTTTAGCGTGGAGCGCCCTTTAAACTTTTCGAGCATTTTTAATGAATGGCAACATGGAAATAATG
AACAAACACTTGAAAAGTTGAAATAATTAAGAAAGAATATTGAAAGCCACAATTATCCTTGCTTTCTTTTACTAGTCATGACCTATAAAAATATATTTTTCTACTATAAAA
AGCCACTGGCACTTTACTTCGGTTTCGGCGGAAGGCAATGGCTTGGAAGTTATAGGCGAAATTAATAAAAATCAACCATTAGACGATTATCCGCTTCTCGTTATTTCTGCGC
GGGGGGTTTTGTGTAGTTTTACTACGCCACTCGGCGGAGCGCGGAAAGTTTTCCCATTTCCTCCATTTGGGTCAGCTGTTGACACAGCGGCCGCCAGAAAACCCATAAAA
AGCGCAAATAAAAATACTTTTAGCGTATTAGTTATGCGTAGACACATCCATCGTTGGTGTATTTAAGATGCTACACTTCCTTTATCAGTTATGGCCACACATTTTCCTGCTC
GAAAAGCTTTTTCATGGTACGAGAATTTCCACCTTTCTACGTGGCACTCTGCTAAACTCGGTAAGACTTGAACAATGGAGCAGTAGCAACTGCAGAATGTATACATTTTCCA
AAAATGTTCCATCAATTTGGAAATGATTTCAGCCAAACAAGAAAATCAAGCAAAAGCACTTTTCACCCATAACCAAGCCGCTTGCCACCCACAATGGAAAGTAGATCAAAGC
CACAACGTTTTGAGTCAGCAGTTTGTTACCATGGACCGCATGTGGGTATTTCCATGTAAACATGCCTGTACTTAGTCCTTGGTGCAATCAGGACATGGATGGTTTGCTAATA
AAAATCTTATTACGGCACCATCTTATTTACACAACACGGTAGAGATATATGGTTTTAAAAATTTCCATTTTCCTGCTTAAAACAGGCAATAAATTGATAGGAAGGTAAATAT
CCTTAAGAAACACTTGTATAAAGTTCAAGTTCAATCCATGTGTCTTAACCTTTAATTGTTTACTATGTTGCATGTTCGTTCCAGTTCTTCATCTTCCATGTAGCGCGTGGCC
CATTCGTGGAGGAGTTTTACCAGTGCGTCACCCACGGATTCTACACGGCGGATTGGCAGGAGCAGATGTACGCCACCTTCACGCTGGTCTTCACCTTCCTGCTGCCGCTGTG
CATCCTGTTTGGCACCTACATGTCCACCTTCCGCACCATTTCCAGTGAGTTTTCAGCGAGAAGGAGTTTACATTCCCGTACTAACCTGATAACCTGATAATCGTGCAGGCAG
CGAAAAGATGTTCCAGGGATCAAAGTTGGCCAACTACTCAACGGCCAAATTGCCCACACAGACGAATCGCCAGAGGCTGATACACAAGGCCAAGATGAAGTCGCTTCGCATA
TCCGTGGTGATCATCATAGCGTTTCTCATCTGCTGGACGCCCTACTACGTCATGATGATTATGTTCATGTTCCTCAATCCGGACAAAAGGGTGAGTTTCCGCCCGTTCCCCT
GCCGAATTTGAAGCTGAATGACAAGCAGATATCGATTCCTGATACCAAAAGGCTCATCAATCAGCAGTCGATTTGTCAGTTTGTAGCGCACGGTTTGGGCGAATCCTTTGTC
CTGTTCCCTACCATATTTACGACTTGGGGTCATGGTTGTACGAGTAAAATAAAAACGAGAATACGACACGACCTAGTCGCAGCTGGCAAATCCGTTTGGAGCTATACCCACT
GCCATAAAAATGATTTCTTATGCTCAATTTTTAATTAATCGTTACGCCACCCGACCGACCACACGGTCGCAATTTTTCCTTAACGGCGTACCCAACTTTGATGACTTGCCAG
AGATGGTCCGTAAAAGCTGGCCCAAGTTGCAAGATTACGGCGAGGCAATCCGGGCAGAAGGAACACAAGGTCGAATTGCTACAAACCGAGCATTCTGTACAGTAGGGGAAAA
TGAATGAGTAGCAAAATATTTCCAAAGTTTATTTGTTTCGATTTAAATTTTAAAAATGCTATAATTGACAAAATAGAATCATTTTTGCTGTCACTATTTTTAAGTGAATGGG
ATATTTGTCCGCCGATTAACCTGAATTTCATATCCTAATGAATTGGATTGCACATTCATAATGGCATTTCCGCTGAAACTCTTTCTCGCTGTGGAAATCCCAGCGATGGCAG
TGTGCAAAATAATTAATTTGGAAGCCAAATTTTTAGCGCAGTGTATCCATAAGCGACCTGAGCCCCAATTCGACCGCAGGAACGCCCACAATGCGCCCACGCGAATGTAGAG
GGGAAAAGGTGAAAGGTGTGCGACAAGGGCAAGGAAATTAATTTGAATTTCACGGATTGCTTGCAGCTGGGCGACGATCTGCAGGACGCCATCTTCTTCTTCGGCATGTCAA
ACAGCCTGGTCAACCCACTCATCTACGGTGCCTTCCACCTGTGTCCTGGCAAAGGGGGCAAGTCGAGCGGCGGGGGCGGCAACAACAACGCCTACAGCTTGAACAGGTGAGT
GGTGAGCCAACAGAATAAGTTAATTCTATTAACCACCACATTTCGTACTTGTTTTTGCGCTTTTATCGTGTCAAAAAATAGTTACTTTTCACTCAAGCAATACACTTTTAAG
ACAGCCAAGCAAGCTTTTGAAGTCTGAATTGCTGCTTTTCTGATCAAAGAGTTTTTATTCTTGGTAATATTTCACATTTAAGAGTATAAATGATAGTGTAAATTTAAATAAA

ATAGTAGAAAGTTTTCCCTAAAGATCGATGTATAACTATTAATACATATTTTACTTAATCTAGTGTTGCTTTTTAATGACAGTTTTTGCTGTGCTTAAATTTTTACAAAATA
AATCTTAAATCCTTGCAAAGAGTTTTCTTTTATGCAAATGGAAAACATTGTTTCATCGCTTGTTAATATCATTTACTTTATTTATTTATTATTCTAAACAGCTGCCAAAGCA
GCGAACTATCTCAGGATCAGATATGATTTGCTATGATTTTGTCACACTTCCTATGTTTGTTTTGATAACGCAAAAAGCACACACACCGTGATTGGCAATAAACTGATAAATT
CTTTTCCGTCTGGTTTTTATCACCTCACCGCATCGTCCATTCCCCCGTCCACTTTCCATCCGCAGGGGCGACTCGCAGCGCACTCCATCCATGCTAACGGCGGTGACGCAGG
TGGACGGCACAGGTGGCAGTTCCCGCCAGATGCGGGCCTTCCGCCAGCAGAGCTACTACCGCAGCTCCTCCAACGGCACAGCCGGACCGGGTGCAGCTCCCTTTAAGGAGCA
GGTTGGCCTGCTGCACGTGGGTCCCGGCAATGGGACGCCCGGTGGCTCCGTCTCTAGCGGCGCCACGCCGCAGTTGATCCGTAAGGGATCGGCTCTGTTGGCCCGACAACCC
AGCTGTCTGAGGGAGCAGGAGCACCAGCAGCGTTTGCTGCTGCACGAGAAGCCCTCGACCCTGGTGCTCAGCTACGACAGCCAGCGGGGCGGAGTGGGCGTGGGCGTGGCCA
GCGGTCTGCTGGACAACAACGAGCGAGTGTCGAGCGTGTGAGAGCACCTGGCGATGGCGGTAGCGGCGGCGGCGGTGGCGGTGGATGCACCGGACACAGTAGAAGCAGCAGC
AGCAACAGCATCAGCCGATGCGGATGCGAATGTGGCGATGGAGGACGCCTGCTCCTGCTGCCATTGCCTGAGCAGGGAGCTGGTGCGGCGGCAGTTGGAGCTCCAAGTGCTT
CCAGTGCAGTGAGGCGACTGCCAGTGGATTTGATCCGGCTGGCTGAGTGGCCGGAGCTGGGGATTCCAGCGATTCCCTCGGTCAACGTGGGGAGTGGCGTCGATGATTAGAT
CACCGGGGCGTGCAACATGTGCTTAGATAGTGCTAATCGAGAGTTAAGTCTTCAAACACAATGGGTGTGCTTAGAAGTAGCGTAGGATTAGAGCTAGTCCGTACGAAGTGGT
TGCCCTACACTGGGAGAAACTCGTGGTCACTTTGCATGTTCTCTGAATGTTCTACTAATACTTGACTTCAACAAGGATACAAATATAAGAGATAATATCTATGTACATCCCA
TAAAACCATAAAAAATTGCAATTAAATCTGGTTGTGAAAAAAATAGATTTAACTTCTGAATGTTTCAAATGTTGCCAATTCGCCAGCGTGTAATCCGATATAAGCAACAATA
ATTACAATTATATAATCAATGTGCAATCAGAGTCAGTTACAATTTACAATTATATGACTGAATAATTGGACTATGAATATATACGACTTCAAATGAGTTTCTCTCATTAGCAGG
GCAACTTTGTAATGTTAATTAACAGCGTTTTTCAGTTTTTCATTTTGTTCCCCCAGCCGAGTAACTCAAAAATCGATTCTATAATTGTGTAAAATAGAAATCATCTAAAAAC
AGGGTAAATCTAACTATTGTAAATACATTTACTCATTATTTTGAGAAGAATATTGTAAAAAAAAAAGATTGTAAAGTTTAAAATAAATATGAGTTCATGTTCAACGACACAA
GGAATTATTTTGATTAATTTAAAGTACTAAAAA
(SEQ ID NO: 196)

Exon: 1001..1210
Exon: 1307..1774
Exon: 6357..6540
Exon: 6605..6810
Exon: 7683..7834
Exon: 8466..8889
Start ATG: 1001

Transcript No. : CT29989
ATGGAGGACGAGTGGGGCTCCTTTGATCGTCTGCCCAGTGTTCCGAGTGCCTCGATGGATTTGGAGACGGAAAACGAGGTGGTCAGCAATTGGTCCACACTGGCCAACTTCA
CGCGACTTGTGGCTGGTGCCGCTCCTGAAATCGTCAACTATACGCTCAACATGATCGACGTGGGTGTGGGCATGGCCACGGATATATCCAATTTGAGCGTTCCGGAGCACGT
GATGGATCACGCACCTCAACTATCCCGATCCGGGTTGCTGAAAGTGTATGTCCTGGCGGTAATGGCACTGTTCTCTCTGCTGGGCAACCTGCTGACCATCTGGAATATCTAC
AAAACCCGCATCTCAAGAAGAAACTCACGGCACACGTGGAGTGCTATCTACTCACTGATGTTCCATCTGTCCATCGCCGATGTCCTGGTCACCTGGTTCTGCATTATCGGGG
AGGCCGCGTGGTGCTACACCGTCCAGTGGCTGGCCAATGAGCTCACCTGCAAGCTGGTGAAGCTCTTCCAGATGTTTAGCCTCTACCTGAGCACCTATGTCCTGGTCCTCAT
CGGAGTGGACCGCTGGATAGCGGTCAAGTATCCGATGAAGTCGCTCAACATGGCCAAGAGGTGTCATAGGCTACTTGGCGGTACTTACATCCTGTCGCTGGTGCTCAGCTTG
CCACAGTTCTTCATCTTCCATGTAGCGCGTGGCCCATTCGTGGAGGAGTTTTACCAGTGCGTCACCCACGGATTCTACACGGCGGATTGGCAGGAGCAGATGTACGCCACCT
TCACGCTGGTCTTCACCTTCCTGCTGCCGCTGTGCATCCTGTTTGGCACCTACATGTCCACCTTCCGCACCATTTCCAGCAGCGAAAAGATGTTCCAGGGATCAAAGTTGGC
CAACTACTCAACGGCCAAATTGCCCACACAGACGAATCGCCAGAGGCTGATACACAAGGCCAAGATGAAGTCGCTTCGCATATCCGTGGTGATCATCATAGCGTTTCTCATC
TGCTGGACGCCCTACTACGTCATGATGATTATGTTCATGTTCCTCAATCCGGACAAAAGGCTGGGCGACGATCTGCAGGACGCCATCTTCTTCTTCGGCATGTCAAACAGCC
TGGTCAACCCACTCATCTACGGTGCCTTCCACCTGTGTCCTGGCAAAGGGGGCAAGTCGAGCGGCGGGGGGCGGCAACAACAACGCCTACAGCTTGAACAGGGGCGACTCGCA
GCGCACTCCATCCATGCTAACGGCGGTGACGCAGGTGGACGGCACAGGTGGCAGTTCCCGCCAGATGCGGGCCTTCCGCCAGCAGAGCTACTACCGCAGCTCCTCCAACGGC
ACAGCCGGACCGGGTGCAGCTCCCTTTAAGGAGCAGGTTGGCCTGCTGCACGTGGGTCCCGGCAATGGGACGCCCGGTGGCTCCGTCTCTAGCGGCGCCACGCCGCAGTTGA
TCCGTAAGGGATCGGCTCTGTTGGCCCGACAACCCAGCTGTCTGAGGGAGCAGGAGCACCAGCAGCGTTTGCTGCTGCACGAGAAGCCCTCGACCCTGGTGCTCAGCTACGA
CAGCCAGCGGGGCGGAGTGGGCGTGGGCGTGGCCAGCGGTCTGCTGGACAACAACGAGCGAGTGTCGAGCGTGTGA
(SEQ ID NO: 197)

Start ATG: 1

MEDEWGSFDRLPSVPSASMDLETENEVVSNWSTLANFTRLVAGAAPEIVNYTLNMIDVGVGMATDISNLSVPEHVMDHAPQLSRSGLLKVYVLAVMALFSLLGNLLTIWNIY
KTRISRRNSRHTWSAIYSLMFHLSIADVLVTWFCIIGEAAWCYTVQWLANELTCKLVKLFQMFSLYLSTYVLVLIGVDRWIAVKYPMKSLNMAKRCHRLLGGTYILSLVLSL
PQFFIFHVARGPFVEEFYQCVTHGFYTADWQEQMYATFTLVFTFLLPLCILFGTYMSTFRTISSSEKMFQGSKLANYSTAKLPTQTNRQRLIHKAKMKSLRISVVIIIAFLI
CWTPYYVMMIMFMFLNPDKRLGDDLQDAIFFFGMSNSLVNPLIYGAFHLCPGKGGKSSGGGGNNNAYSLNRGDSQRTPSMLTAVTQVDGTGGSSRQMRAFRQQSYYRSSSNG
TAGPGAAPFKEQVGLLHVGPGNGTPGGSVSSGATPQLIRKGSALLARQPSCLREQEHQQRLLLHEKPSTLVLSYDSQRGGVGVGVASGLLDNNERVSSV*
(SEQ ID NO: 198)

Name: GONADOTROPIN-RELEASING HORMONE RECEPTOR-like
Classification: G_protein_linked_receptor

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/09341

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/435 C07K14/705 C07K16/18 C12N5/10
C12Q1/68 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K G01N C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

SEQUENCE SEARCH, EPO-Internal, BIOSIS, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL SEQUENCE LIBRARY [Online] 6 January 2000 (2000-01-06) ADAMS, M. AND VENTER, J.C.: "Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces" XP002182628 accession no AC020076 ---	1,2,4,5
A	WO 94 08006 A (ZYMOGENETICS INC) 14 April 1994 (1994-04-14) the whole document ---	

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 November 2001	15. 02. 2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LI X-J ET AL: "CLONING, FUNCTIONAL EXPRESSION, AND DEVELOPMENTAL REGULATION OF A NEUROPEPTIDE Y RECEPTOR FROM DROSOPHILA MELANOGASTER" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 267, no. 1, 5 January 1992 (1992-01-05), pages 9-12, XP000877443 ISSN: 0021-9258 the whole document ---	
A	WO 99 01468 A (DEN HEUVEL MARCEL VAN ;INGHAM PHILIP W (GB); ONTOGENY INC (US)) 14 January 1999 (1999-01-14) the whole document ---	
A	HAUSER FRANK ET AL: "Molecular cloning, genomic organization and developmental regulation of a novel receptor from Drosophila melanogaster structurally related to gonadotropin-releasing hormone receptors from vertebrates." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 249, no. 3, 28 August 1998 (1998-08-28), pages 822-828, XP002182627 ISSN: 0006-291X the whole document ---	
A	FENG G ET AL: "CLONING AND FUNCTIONAL CHARACTERIZATION OF A NOVEL DOPAMINE RECEPTOR FROM DROSOPHILA MELANOGASTER" JOURNAL OF NEUROSCIENCE, NEW YORK, NY, US, vol. 15, no. 12, 15 June 1995 (1995-06-15), pages 3925-3933, XP002919142 ISSN: 0270-6474 the whole document ---	
P,X	ADAMS M D ET AL: "THE GENOME SEQUENCE OF DROSOPHILA MELANOGASTER" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 287, no. 5461, 24 March 2000 (2000-03-24), pages 2185-2195, XP000961051 ISSN: 0036-8075 the whole document ---	1,2,4,5

-/--

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/09341

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-20 partially

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: claims 1-20 partially

Isolated protein consisting or comprising an amino acid sequence as characterized by SEQID3, or an allelic variant or an ortholog of said amino acid sequence wherein said variant or ortholog is encoded by a nucleic acid molecule that hybridizes to the nucleic acid molecule as characterized by SEQID1 or 2; an antibody that binds to said protein; furthermore a nucleic acid molecule consisting or comprising of a nucleotide sequence that
1) encodes the amino acid sequence of SEQId3
2) encodes and allelic variant or an ortholog of an amino acid sequence of SEQId3 wherein said nucleotide sequence hybridizes to SEQID 1 or 2
3) encodes a fragment of said SEQID3
4) is the complement of the nucleotides of 1) to 3)

The recombinant expression of the same in host cells and methods for the detection of said proteins or said nucleic acids in a sample with the help of an agent that binds to said protein or an oligonucleotide and kits that contain such agent or oligonucleotide.
Furthermore, a method to identify an agent that binds to said protein by detecting a complex formed by an agent and the said protein.

Invention 2-66: claims 1-20 partially

as invention one but referring to the protein and nucleic acid sequences as characterized by SEQIDs 6,9,...,192,195,198; SEQIDs 4,7,....,190,193,196 and SEQIDs 5,8,....,191,194,197, respectively.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/09341

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE EMBL SEQUENCE LIBRARY [Online] 24 March 2000 (2000-03-24) ADAMS, M.D:, ETAL.: "Drosophila melanogaster genomic scaffold 142000013386053 section 8 of 30 complete sequence" XP002182629 accession no. AE003491 and AE002593 ---	1,2,4,5
P,X	DATABASE TREMBLREL. DATABASE [Online] 1 May 2000 (2000-05-01) ADAMS, M.D., ET AL.: "The genome sequence of Drosophila melanogaster - cross-reference to EMBL accession no. AE003491, AAF48216.1" XP002182630 accession no. Q9VYH9 ---	1,2
E	EP 1 136 501 A (BAYER AG) 26 September 2001 (2001-09-26) the whole document -----	1-20

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/09341

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 9408006	A	14-04-1994	CA	2145866	A1	14-04-1994
			EP	0663006	A1	19-07-1995
			JP	8501942	T	05-03-1996
			WO	9408006	A1	14-04-1994
			US	5683884	A	04-11-1997
			US	5674981	A	07-10-1997
			US	5622839	A	22-04-1997
			US	5674689	A	07-10-1997
WO 9901468	A	14-01-1999	AU	8380898	A	25-01-1999
			WO	9901468	A2	14-01-1999
EP 1136501	A	26-09-2001	DE	10013618	A1	20-09-2001
			EP	1136501	A2	26-09-2001
			JP	2001299369	A	30-10-2001